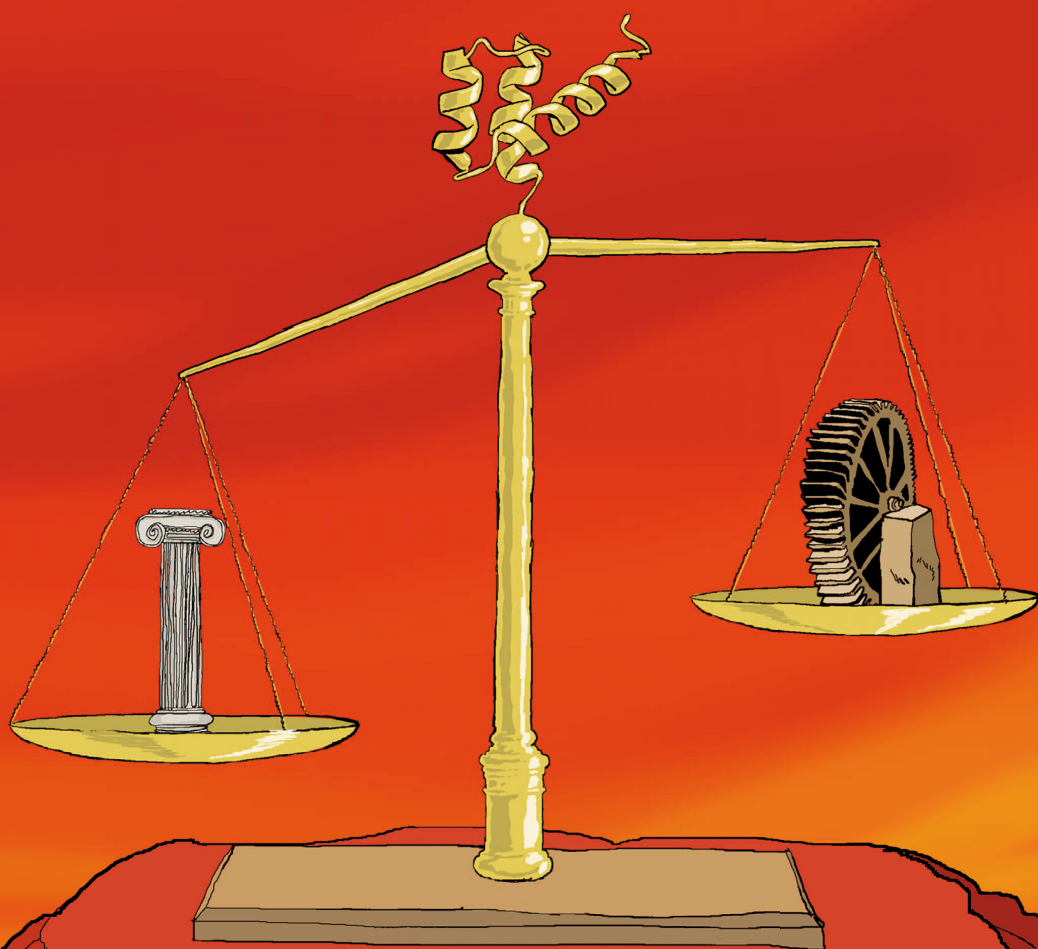


ESTUDIO TERMODINÁMICO Y
ESTRUCTURAL DEL MOTIVO DE
UNIÓN A ADN DE HESX-1:
RELACIÓN ESTABILIDAD-FUNCIÓN
EN HOMEODOMINIOS



MARIO TORRADO DEL REY

Mario Torrado del Rey

Estudio termodinámico y estructural del
motivo de unión a ADN de HESX-1:
relación estabilidad-función en homeodominios



DEPARTAMENTO DE BIOLOGÍA MOLECULAR
FACULTAD DE CIENCIAS
UNIVERSIDAD AUTÓNOMA DE MADRID

**ESTUDIO TERMODINÁMICO Y ESTRUCTURAL DEL
MOTIVO DE UNIÓN A ADN DE HESX-1:
RELACIÓN ESTABILIDAD-FUNCIÓN EN HOMEODOMINIOS**

MEMORIA
que para optar al grado de Doctor
presenta

Mario Torrado del Rey

Madrid, 2012

Director: Dr. Juan Luis Asensio Álvarez

Tutor: Dr. José María Carazo García

Departamento de Química Bio-Orgánica
Instituto de Química Orgánica General
Consejo Superior de Investigaciones Científicas



Madrid, julio 2012

La presente tesis ha sido realizada en el
Departamento de Química Bio-Orgánica del
Instituto de Química Orgánica General del
Consejo Superior de Investigaciones Científicas,
bajo la dirección del Dr. Juan Luis Asensio.

Para ello, Mario Torrado ha contado con la ayuda de una
beca/contrato de Formación de Profesorado Universitario (FPU),
que le fue concedida por el Ministerio de Educación, Cultura y Deporte.

El trabajo ha sido financiado por el “Plan Nacional de I+D+i”
(Ministerio de Economía y Competitividad) a través de los proyectos
CTQ2007-67403/BQU y CTQ2004-04994/BQU.

Parte de los resultados obtenidos en este trabajo
se recogen en el siguiente artículo:

Torrado, M., Revuelta, J., Gonzalez, C., Corzana, F., Bastida, A., Asensio, J.L., 2009.
Role of conserved salt bridges in homeodomain stability and DNA binding.
The Journal of Biological Chemistry **284**, 23765-23779.

Ilustración de la cubierta: Carlos Vermut.
www.pscosoda.com

Fotografía de la cubierta cortesía de Juan Navalón e Imágenes y Sensaciones S.L.

Prólogo de Florián Reyes Sampablo

Dedicado a mis padres y hermanos, autores de honor de esta tesis

ÍNDICE

<i>PRÓLOGO</i>	xiii
<i>ABSTRACT (RESUMEN EN INGLÉS)</i>	xix
1. INTRODUCCIÓN	1
1.1. EL HOMEODOMINIO	3
<i>1.1.1. Estructura del homeodominio</i>	3
<i>1.1.2. Interacciones terciarias conservadas en homeodominios</i>	6
<i>1.1.3. Estabilidad de los homeodominios</i>	11
1.2. UNIÓN DEL HOMEODOMINIO A ADN	19
<i>1.2.1. Secuencia de reconocimiento específico de homeodominios</i>	19
<i>1.2.2. Características estructurales de los complejos específicos</i> <i>homeodominio/ADN</i>	20
<i>1.2.3. Interacciones observadas en los complejos específicos</i> <i>homeodominio/ADN</i>	22
<i>1.2.4. Termodinámica de unión de homeodominios a ADN</i>	26
<i>1.2.5. Unión inespecífica de homeodominios. Relevancia funcional</i>	32
1.3. PATOLOGÍAS HUMANAS RELACIONADAS CON LA DISFUNCIÓN DE HOMEODOMINIOS	35
2. ANTECEDENTES Y OBJETIVOS	39
2.1. ANTECEDENTES DEL GRUPO DE INVESTIGACIÓN: EL HOMEODO- MINIO DEL FACTOR DE TRANSCRIPCIÓN HUMANO HESX-1	41
2.2. PLANTEAMIENTO DEL OBJETIVO PRINCIPAL: LA RELACIÓN ENTRE LA ESTABILIDAD Y LA FUNCIÓN DE HOMEODOMINIOS	43
2.3. DESARROLLO DEL OBJETIVO PRINCIPAL Y DE LOS OBJETIVOS SECUNDARIOS	48
<i>2.3.1. Objetivo principal</i>	48
<i>2.3.2. Objetivos secundarios</i>	51
<i>2.3.3. Procedimientos experimentales destinados al abordaje de los</i> <i>objetivos</i>	52

3. RESULTADOS	53
3.1. ANÁLISIS TERMODINÁMICO Y ESTRUCTURAL DE LAS VARIANTES DEL HOMEODOMINIO DE HESX-1.....	55
3.1.1. <i>Parámetros termodinámicos del homeodominio de HESX-1 salvaje.....</i>	55
3.1.2. <i>Análisis de secuencias de homeodominios y diseño de los mutantes.....</i>	55
3.1.3. <i>Análisis de la estabilidad de las variantes de HESX-1: aspectos metodológicos.....</i>	62
3.1.4. <i>Puentes salinos frente a pares hidrófobos en las posiciones 19-30.....</i>	66
3.1.5. <i>Interacción entre el par 19-30 y el residuo 26 del lazo I-II: influencia en la estabilidad del homeodominio.....</i>	70
3.1.6. <i>Puentes salinos frente a pares hidrófobos en las posiciones 31-42.....</i>	74
3.1.7. <i>Determinación de la estructura 3D en disolución de la va- riante termoestable 31 (mutante R31L/E42L) mediante RMN.....</i>	78
3.1.8. <i>Puentes salinos frente a pares hidrófobos en las posiciones 17-52.....</i>	90
3.1.9. <i>Influencia del empaquetamiento de las cadenas laterales en el efecto estabilizador de los puentes salinos.....</i>	91
3.1.10. <i>Análisis de valores pK_a en una selección de mutantes múlti- ples por RMN.....</i>	94
3.1.11. <i>Puentes salinos frente a pares hidrófobos como determi- nantes de la estabilidad del homeodominio: resumen de resultados e influencia del grado de exposición de los residuos involucrados.....</i>	99
3.2. ANÁLISIS DE LAS PROPIEDADES DE UNIÓN A ADN DE UNA SELECCIÓN DE VARIANTES DEL HOMEODOMINIO.....	103
3.2.1. <i>Propiedades de unión a ADN de las variantes con sustituciones en posiciones 31 y 42. Importancia del contacto R31-fosfato.....</i>	104
3.2.2. <i>Propiedades de unión a ADN de las variantes con sustituciones en las posiciones 19, 30, 15 y 37. Relación estabilidad-función en el homeodominio.....</i>	107
4. DISCUSIÓN	115
4.1. ESTUDIO DE LA ESTABILIDAD DE LAS VARIANTES DEL HOMEODOMINIO DE HESX-1: PUENTES SALINOS FRENTE A PARES HIDRÓFOBOS	119

4.1.1. Los puentes salinos en posiciones 19-30 y 31-42 resultan desestabilizantes respecto a pares hidrófobos, mientras que el efecto contrario se observa para el par 17-52	119
4.1.2. Discusión global sobre la aportación a la estabilidad del homeodominio de los puentes salinos respecto a pares hidrófobos. Influencia del grado de exposición al disolvente de los residuos involucrados	122
4.1.3. Funciones de los puentes salinos que justifican su alta frecuencia de aparición en proteínas naturales	125
4.1.4. Correlación entre la posición 26 y la naturaleza del par 19-30: influencia de la tríada 19-26-30 en la estabilidad del homeodominio	127
4.1.5. Los metilenos de las cadenas laterales involucradas en puentes salinos contribuyen significativamente a la estabilidad del plegamiento	128
4.2. ESTUDIO DE LAS PROPIEDADES DE UNIÓN A ADN DE UNA SELECCIÓN DE VARIANTES DE HESX-1: RELACIÓN ENTRE ESTABILIDAD Y FUNCIÓN EN HOMEODOMINIOS Y SU IMPORTANCIA EVOLUTIVA.....	130
4.2.1. El puente salino 31-42, altamente conservado en homeodominios, tiene un papel modesto en el reconocimiento de ADN	130
4.2.2. Las propiedades de unión a ADN de las variantes con sustituciones en las posiciones 19, 30, 15 y 37 revelan una independencia entre la estabilidad y la función del homeodominio	132
5. PROCEDIMIENTOS EXPERIMENTALES	135
5.1. OBTENCIÓN DE LAS VARIANTES DEL HOMEODOMINIO DE HESX-1.....	137
5.1.1. Construcción de plásmidos y mutagénesis dirigida	137
5.1.2. Expresión y purificación proteica	143
5.2. EXPERIMENTOS DE DICROÍSMO CIRCULAR.....	147
5.2.1. Curvas de desnaturalización química	147
5.2.2. Curvas de desnaturalización térmica	148
5.3. EXPERIMENTOS DE RESONANCIA MAGNÉTICA NUCLEAR (RMN).....	153
5.3.1. Experimentos de RMN monodimensionales de los homeodominios salvaje, 31, 17 y 19	153
5.3.2. Determinación de la estructura tridimensional de la variante termoestable 31 (R31L/E42L) por RMN	153
5.3.3. Medidas de pK_a en los mutantes cuádruples 12 y 15 por RMN	155
5.4. EXPERIMENTOS DE UNIÓN A ADN POR MICROCALORIMETRÍA.....	157

CONCLUSIONES..... 160

REFERENCIAS..... 162

LISTA DE ABREVIATURAS..... 182

AGRADECIMIENTOS 186

PRÓLOGO

Quisiera ser homeodominio, pero solo a veces

Los homeodominios no existen. No, no existen al menos a los ojos del gran público, de la gente de a pie, en su mundo y en el mío... Pero para aquellos que saben mirar, como Mario, sí que existen y resulta que son maravillosos.

Servidor, que es verdadero profano en bioquímica teórica, que no de la práctica pues aprecio sobremanera la fermentación en sus más amplias aplicaciones, reconozco que escribir el prólogo de esta tesis ha sido todo un reto. Cómo explicar, plasmar en unas pocas líneas, y sin desvelar secreto alguno ni ahuyentar al lector para que se adentre en la aventura que supone esta trabajada tesis, qué es un homeodominio y cómo funciona pero no para los doctos en la materia, a los que arrancaré una sonrisa a lo sumo, sino a mis ojos, a los del gran público, a los de la gente de a pie.

Y nada mejor que la relativización de un concepto fuera de su entorno para intentar explicarlo. Muchas fueron las vueltas, se lo aseguro, las que le di buscando los parámetros y las variables para alejar a HESX-1 de su atareada función de unirse al ADN para regular la expresión de ciertos genes, sacarlo de la célula y, como en la películas de ciencia ficción, que atravesase el telón para poder así tocarlo casi con los dedos.

Por lo general, la introducción de mutaciones en las proteínas genera una alteración en su estabilidad, y con ello normalmente un cambio en su funcionalidad, lo que puede llegar a acarrear defectos en el organismo y en nuestro caso, el caso de las personas, numerosas enfermedades.

He aquí que me doy cuenta que bajo este prisma y extrapolando las variables, nuestro funcionamiento social, el humano, no se aleja demasiado del de una proteína cualquiera. Los acontecimientos que nos van ocurriendo a lo largo de nuestra vida mutan aunque no queramos nuestra realidad y, sin duda, alteran nuestra estabilidad (unas veces a mejor y otras a peor) con la consecuente variación en nuestra manera de actuar, o lo que básicamente es lo mismo, en nuestra funcionalidad.

¿Cuántos no han cambiado su vida y su trabajo al tener un hijo o quién no ha mentido por amor cuando nunca lo había hecho? Y sin ir más lejos, el

autor de esta tesis, al que le ha costado más tiempo del previsto terminarla, pero no por falta de ganas sino por algunas de esas mutaciones ocurridas en su vida al terminar la carrera y encontrar, por ejemplo y entre otras cosas, todo un mundo de variables nuevas y apasionantes fuera de esas aulas y laboratorios que tantas horas, días y meses fueron su espacio natural.

Llegados a este punto, Mario me comentaba que hasta aquí todo esto le parecía muy correcto, que estaba de acuerdo con mi punto de vista, pero que su proteína, HESX-1, no es como las demás, es diferente. Porque lo apasionante de esta tesis, ya lo verá usted amigo lector, es que con algunas mutaciones realizadas al homeodominio de HESX-1, el nexa que debería existir entre su estabilidad y su función no resulta aparente. En palabras del autor, “una de las cosas bonitas que hemos visto es que algunas de las mutaciones que hemos introducido afectan a la estabilidad del homeodominio de HESX-1, pero sin afectar para nada a su función”. Y resulta que Mario y yo tenemos una persona en común que es tan especial como su HESX-1... Mi hermano Gerardo.

Gerardo es amigo de los hermanos Torrado desde hace muchos años y hace también muchos años que tuvo un accidente de coche que le dejó como secuela un daño cerebral. Básicamente, mi hermano tiene lo que los médicos llaman un daño axonal difuso en la zona más básica y primitiva del cerebro, lo que es lo mismo, si me permite usted que lo simplifique, que decir que no saben exactamente lo que tiene pero que lo tiene pues su día a día, y yo se lo aseguro, así lo certifica. Pero el cuerpo humano siempre encuentra lo que mi padre denominaría “el bypass correcto” para minimizar los daños e incluso potenciar otras facetas tal y como normalmente le ocurre a un ciego con su sentido del oído y el olfato.

Pero Gerardo es diferente. Como en el caso de HESX-1, aunque algunas sí lo hacen, muchas de esas mutaciones o acontecimientos a los que antes me refería no le afectan en su función. Es raro que mi hermano no duerma bien por la noche por duro que haya sido el día y se levante, muy temprano, todas las mañanas, con todas sus ganas para empezar el día. Aunque su estado de ánimo pueda no ser el mejor, su capacidad de trabajo, muy artesanal para mayor dificultad, no se ve alterada y sigue siendo igual de cuidadoso y metódico. Su cuerpo ha elevado su umbral del dolor muy por encima de la media y si Gerardo se cae, y lo hace a menudo, el golpe que a todos nos dejaría en casa postrados y doloridos durante días y alteraría claramente nuestra funcionalidad, apenas le supone algún mínimo y lógico ajuste ese día.

Claramente, el equilibrio y nexa de unión entre estabilidad y función, no siempre se refleja claro. Ni mucho menos. Yo hasta ahora no me lo había planteado pero ahora veo que tiene más lógica de lo que parece y que HESX-1 y sus iguales no dejan de comportarse como a mí me gustaría comportarme muchas veces. Supongo que es eso que siempre se ha dicho de tener la ‘sangre

fría' y no dejar que ciertas cosas nos afecten tanto que alteren nuestro día a día, nuestra función. Pero soy de los que piensa que la vida no se escoge sino que la vida se vive y hacerlo conlleva mutar con ella y descubrir como la senda va cambiando con nosotros, algo que los homeodominios, según me sugiere Mario, no pueden hacer.

Florián Reyes Robles, in memoriam.

FLORIÁN REYES SAMPABLO

This thorn in my side...
This thorn in my side is from the tree.
This thorn in my side is from the tree I have planted.

JAMES HETFIELD, 1996
"Bleeding Me"

ABSTRACT
(RESUMEN EN INGLÉS)

Homeodomain proteins are transcription factors present in all eukaryotes and play key roles in cellular differentiation during development. In fact, mutations in homeodomain-encoding genes give rise to several inherited diseases in humans. The homeodomain consists of a 60-residue DNA binding domain composed of disordered N- and C-terminal regions and three helical segments. These structural features, together with the highly charged nature of the polypeptide, determine a typically low stability, property also shared by other DNA-binding proteins. Conversely, homeodomains usually present a large DNA-binding affinity. From a molecular recognition perspective, the association process is sequence-specific and represents an example of the “induced fit” mechanism. Thus, it has been shown that disordered regions of the protein become structured upon complex formation, which highlights the relevance of the protein flexibility in this process. Herein we have investigated the relations between stability and DNA-binding in the human HESX-1 homeodomain.

The sequence information available for homeodomains reveals that salt bridges connecting pairs 19-30, 31-42, and 17-52 are frequent, whereas aliphatic residues at these sites are rare and mainly restricted to proteins from homeotherms. We have analyzed the influence of salt and hydrophobic bridges at these sites on the stability and DNA-binding properties of human HESX-1 homeodomain. Regarding the protein stability, our analysis shows that hydrophobic side chains are clearly preferred at positions 19-30 and 31-42. This stabilizing influence results from the more favourable packing of the aliphatic side chains with the protein core, as illustrated by the three-dimensional solution structure of a thermostable variant, herein reported. In contrast, only polar side chains seem to be tolerated at positions 17-52. Interestingly, despite the significant influence of pairs 19-30 and 31-42 on the stability of the homeodomain, their effect on DNA binding ranges from modest to negligible. In other words, this analysis reveals an intriguing lack of correlation between binding strength and the homeodomain conformational stability. Previous studies have shown that the induced fit of the protein disordered regions plays

a key role in the molecular recognition process. On the contrary, according to our data, the potential adjustment of the structured core of the protein seems to have little influence on the complex stability. This features might allow and independent modulation of stability and DNA-binding by evolution.

1 **INTRODUCCIÓN**

1.1

EL HOMEODOMINIO

El homeodominio es un motivo estructural de reconocimiento de ADN altamente conservado en organismos eucariotas, y presente en un gran número de factores de transcripción que reciben el nombre de homeoproteínas. El segmento de ADN que codifica el homeodominio se conoce como homeobox, y fue inicialmente identificado como un rasgo común presente en muchos de los genes que especifican la información posicional y temporal en las rutas reguladoras del desarrollo embrionario de *Drosophila*. Hasta ahora han sido descritos cientos de genes homeobox en multitud de especies eucariotas, muchos de los cuales intervienen en procesos clave del desarrollo, así como en el mantenimiento de la homeostasis (Wolberger, 1996; Stollar *et al.*, 2003). En todos los casos, la acción de estos factores de transcripción se inicia mediante la interacción del homeodominio con secuencias de ADN presentes en regiones reguladoras de sus genes diana, promoviendo o reprimiendo la transcripción de los mismos. Dicho proceso de reconocimiento molecular se caracteriza por una elevada afinidad y una exquisita especificidad de secuencia. De hecho, este último rasgo de la asociación es el que define, en última instancia, la función biológica del factor de transcripción (Ekker *et al.*, 1992). En general, aquellas mutaciones que afectan de forma significativa a la capacidad de los homeodominios de ejercer su función suelen tener graves consecuencias sobre el desarrollo, produciendo, con frecuencia, severas anormalidades. Esta observación resalta la vital importancia de estos motivos estructurales.

1.1.1. Estructura del homeodominio

A pesar de estar presentes en todo el abanico de organismos eucariotas, las secuencias de los diferentes homeodominios muestran un grado de conservación sorprendentemente alto. Ello es indicativo de una elevada presión evolutiva en este sentido. Como consecuencia, la estructura tridimensional de

estos motivos está, a su vez, muy conservada (Gehring *et al.*, 1994a). Sus rasgos más característicos se describen a continuación.

El homeodominio consta de 60 aminoácidos, y se pliega en tres hélices α (a las que nos referiremos en adelante como I, II y III) que empaquetan dando lugar a un núcleo hidrófobo compacto, tal y como se ilustra en la [Figura 1.1](#) (Wolberger, 1996). Las hélices I (residuos 10 a 22) y II (residuos 28 a 38) son antiparalelas, y están conectadas por un lazo flexible (al que nos referiremos en adelante como lazo I-II). La hélice III (residuos 42 a 59), forma un ángulo aproximado de 60° con las otras dos y es más larga que ellas. Se conecta a la hélice II a través de un giro β de tres residuos (denominado giro II-III), lo cual da lugar a un motivo conocido como hélice-giro-hélice (*helix-turn-helix*), descrito inicialmente en varias proteínas reguladoras procarióticas (Ohlendorf *et al.*, 1982; Otting *et al.*, 1988; Gehring *et al.*, 1994a). Las hélices I y II presentan, en general, un carácter marcadamente anfipático y contribuyen de forma sustancial al núcleo hidrófobo del homeodominio. Por su parte, la hélice III incluye regiones claramente diferenciadas, lo que ha llevado a ciertos autores a dividirla en dos elementos estructurales diferentes (hélices III y IV) (Otting *et al.*, 1990). En adelante, cuando sea preciso resaltar esta división nos referiremos a dichas porciones como hélices III-a y III-b. La primera (residuos 42 a 52) es anfipática y está perfectamente estructurada en el homeodominio

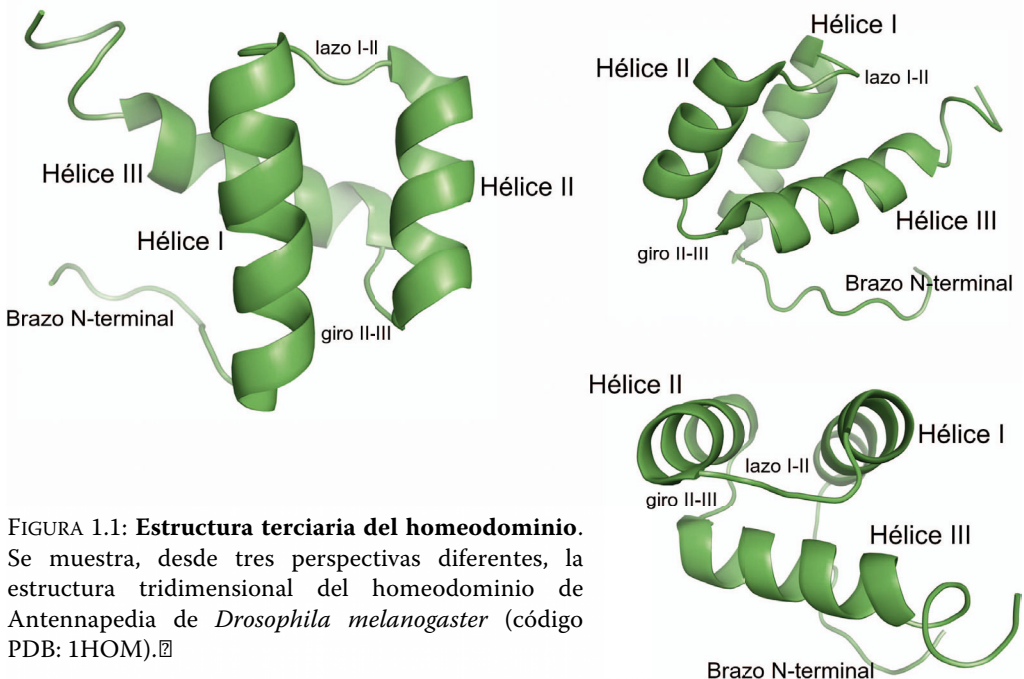


FIGURA 1.1: **Estructura terciaria del homeodominio.** Se muestra, desde tres perspectivas diferentes, la estructura tridimensional del homeodominio de Antennapedia de *Drosophila melanogaster* (código PDB: 1HOM).²

libre. La segunda (residuos 53 a 59) es fundamentalmente polar e incluye un buen número de residuos con carga positiva (argininas y/o lisinas).

Diversos estudios de RMN han demostrado que el homeodominio presenta, en su estado libre, dos zonas parcialmente desordenadas (Güntert

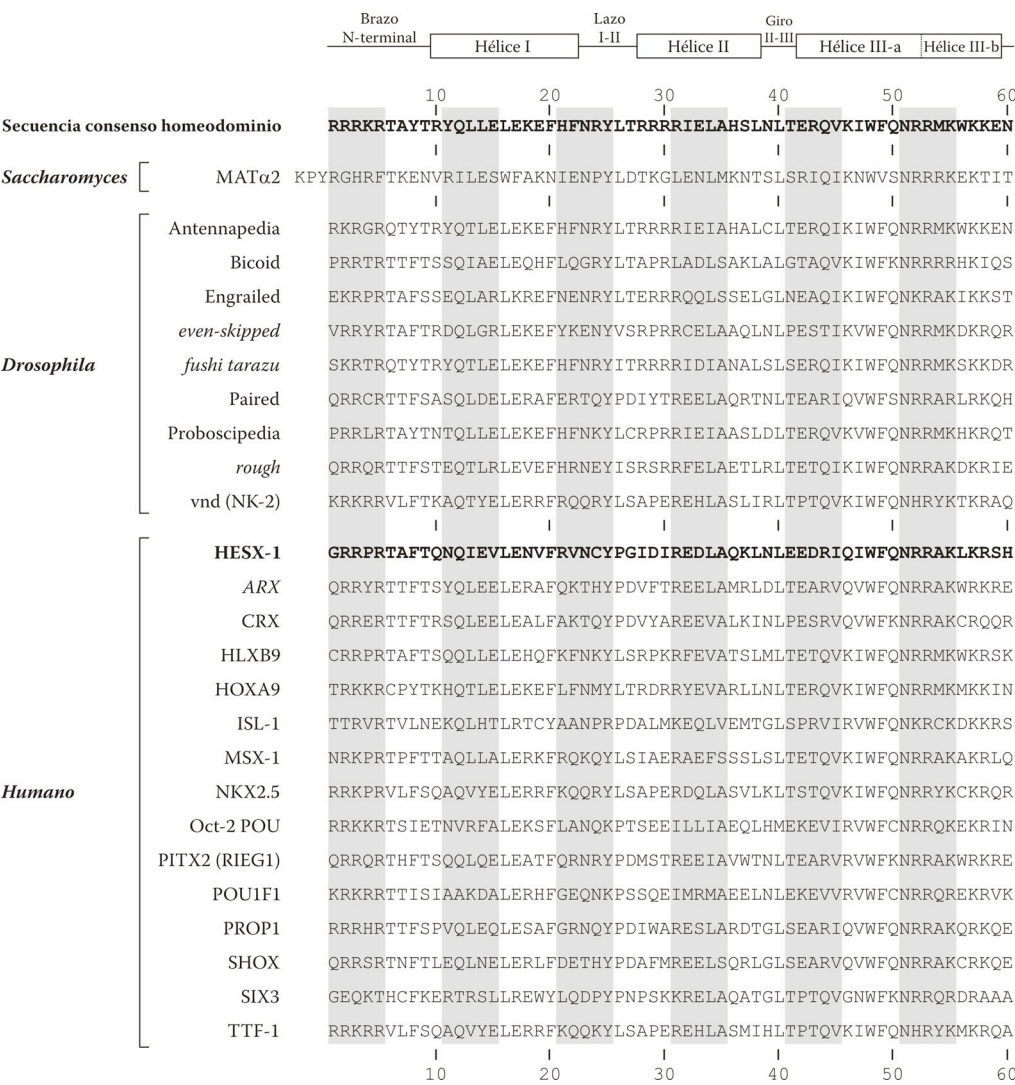


FIGURA 1.2: **Estructura primaria y secundaria del homeodominio.** En la parte superior se muestra la distribución de elementos de estructura secundaria en los homeodominios a lo largo de su secuencia de 60 residuos. En la parte inferior se expone la secuencia consenso del homeodominio (Gehring *et al.*, 1994b), alineada con las secuencias de varios homeodominios representativos (Banerjee-Basu *et al.*, 2000, 2001, 2003), incluyendo el de HESX-1, del que trata esta tesis.

et al., 1991; Phillips *et al.*, 1991; Qian *et al.*, 1994a; Cox *et al.*, 1995; Tsao *et al.*, 1995; Ippel *et al.*, 1998) (ver Figuras 1.3c y 1.7). La primera se denomina comúnmente brazo N-terminal e incluye los residuos 1 a 8. La segunda comprende la hélice III-b previamente descrita, y se encuentra desordenada o en forma helicoidal más o menos definida, dependiendo del homeodominio. Curiosamente, ambas regiones se caracterizan por su carácter policatiónico y juegan un papel clave en la formación del complejo proteína/ADN (ver sección 1.2). Sin embargo, la eliminación de una u otra por mutagénesis no produce cambios destacables en el plegamiento global del homeodominio (Clarke *et al.*, 1994; Qian *et al.*, 1994b; Shah *et al.*, 2007).

La secuencia consenso del homeodominio, así como su distribución de elementos de estructura secundaria, se representan esquemáticamente en la Figura 1.2.

1.1.2. Interacciones terciarias conservadas en homeodominios

Una comparación estructural de diversos homeodominios revela la existencia de ciertos patrones de interacción altamente conservados y que, presumiblemente, juegan un papel clave en la estabilización del plegamiento.

En primer lugar, el núcleo hidrófobo se constituye, en todos los casos, en torno a un empaquetamiento definido de residuos aromáticos, centrado en el triptófano 48 y la fenilalanina 49 de la hélice III (que adoptan una geometría tipo T), y que incluye también las posiciones 8 y 20 (típicamente fenilalanina o, en menor proporción, tirosina) (Figura 1.3). El papel clave de estas posiciones en la estabilización del plegamiento ha sido demostrado mediante el análisis de diversos mutantes simples. Así pues, la sustitución F49I en el homeodominio de *rough* de *Drosophila*, causa un fenotipo de pérdida de función (Heberlein *et al.*, 1994). De un modo similar, las mutaciones F49A y W48A en HESX-1, llevadas a cabo por nuestro grupo de investigación (resultados no publicados), conllevan una pérdida sustancial de estructura secundaria y un descenso de T_m de unos 30 °C. Curiosamente, el apilamiento entre las posiciones 48 y 49 parece ser tremendamente sensible a la naturaleza y tamaño de los sistemas aromáticos involucrados. Así, se ha observado que sustituciones más conservadoras, tales como W48F en Bicoid, causan igualmente una caída significativa en la estabilidad térmica del polipéptido (reducción en T_m de 17 °C) (Subramaniam *et al.*, 2001). Finalmente, de acuerdo con este último trabajo, la mutación F8A tiene como consecuencia un descenso menor pero sustancial de estabilidad (reducción en T_m de 7 °C).

Las hélices I y II, por su parte, contribuyen al empaquetamiento mediante residuos alifáticos, destacando el papel de L34 y L16 (Figura 1.3, a y b). Este

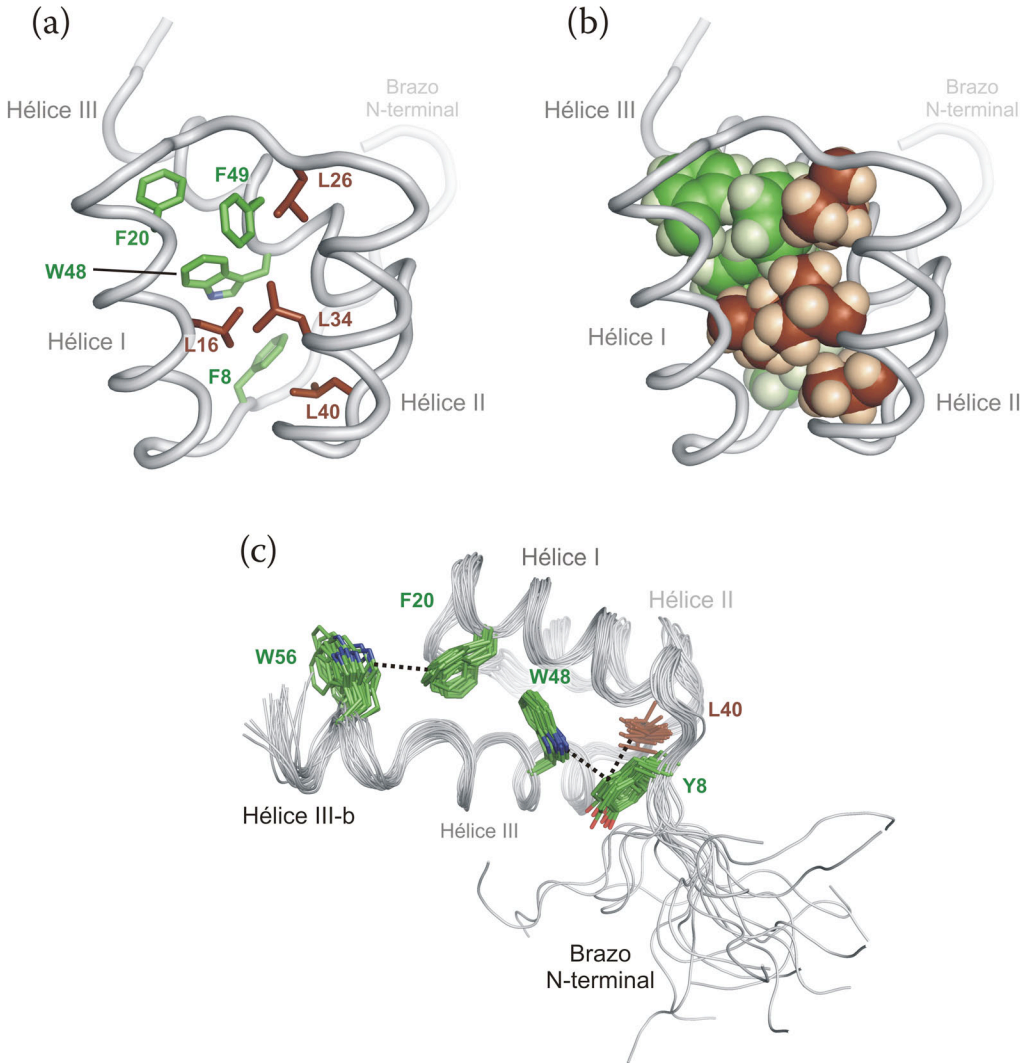


FIGURA 1.3: **Interacciones conservadas del núcleo hidrófobo de los homeodominios.** Los residuos aromáticos se muestran en verde y los alifáticos en marrón. (a) y (b) Empaquetamiento de los residuos aromáticos F8, F20, W48 y F49, y de las leucinas 16, 26, 34 y 40 en el homeodominio de Engrailed (código PDB: 2JW1) en representación de (a) varillas y (b) esferas. (c) Puntos de anclaje proporcionados por el núcleo hidrófobo a las regiones desordenadas del homeodominio de Antennapedia (código PDB: 1HOM). La interacción W56-F20 favorece la estructuración de la hélice III-b, y las interacciones Y8-W48 e Y8-L40 sujetan el brazo N-terminal.

último residuo está conservado en la práctica totalidad de las secuencias de homeodominio descritas. No resulta sorprendente, por lo tanto, su relevancia en la estabilización de estos sistemas. A modo de ejemplo, se ha demostrado

que la mutación simple L16A en Engrailed causa una pérdida de estructura terciaria (claramente reflejada en los espectros de RMN) así como una desestabilización de 4 kcal/mol (Mayor *et al.*, 2003).

Por lo que respecta al lazo I-II, éste contribuye a completar el núcleo hidrófobo de la proteína a través de la posición 26 (prolina o aminoácido alifático en la inmensa mayoría de los homeodominios). Ésta parece contribuir a la estabilidad del plegamiento por medio de contactos con el anillo aromático de F49 y la cadena lateral de L34, si bien su relevancia no ha sido demostrada experimentalmente. De forma similar, la leucina en posición 40 del giro II-III, también extremadamente conservada, interacciona de forma laxa con L16 y el anillo bencénico localizado en la posición 8 (Figura 1.3).

Es de destacar que el núcleo hidrófobo del homeodominio ofrece también puntos de anclaje para las regiones desordenadas de la estructura (Figura 1.3c). Así, los contactos previamente mencionados de la posición 8 con W48 y L40, limitan la libertad conformacional del brazo N-terminal en esa zona. De un modo similar, el residuo F20 puede establecer interacciones con la posición 56 de la hélice III, en aquellos casos en que dicha posición está ocupada por residuos hidrófobos. Este contacto supone un punto de apoyo adicional para la hélice III-b que facilita su estructuración. Su relevancia ha sido demostrada mediante mutaciones estabilizantes como T56W en vnd (NK-2) (Weiler *et al.*, 1998; Gonzalez *et al.*, 2001), así como por un análisis comparado de estructuras de homeodominios (Ippel *et al.*, 1998).

Por lo que respecta a las regiones expuestas del homeodominio, lógicamente su grado de conservación es inferior al observado en el núcleo hidrófobo, excepto en aquellas regiones de relevancia funcional, es decir, involucradas en el reconocimiento de ADN. Sin embargo, una comparación de secuencias pone de manifiesto varias tendencias generales. Así, se aprecia un predominio general de residuos polares (neutros o cargados). En regiones de hélice α es relativamente frecuente la conexión, mediante puentes salinos expuestos, de las posiciones i , $i+3$ e i , $i+4$, especialmente en las hélices I y II. De un modo similar, destaca la posibilidad de formación de diversos puentes salinos entre segmentos diferentes de hélice α . Dos de ellos, altamente conservados en homeodominios, conectan las hélices I-III y II-III, a través de las posiciones 17-52 y 31-42, respectivamente (Figura 1.4, b y c). La importancia del primero ha sido puesta de manifiesto por estudios de mutagénesis, de acuerdo con los cuales su introducción en vnd (NK-2) (mediante la sustitución H52R [Weiler *et al.*, 1998; Gonzalez *et al.*, 2001]) o en Engrailed (mediante la mutación K52E, que hay que tener en cuenta que además supone eliminar la repulsión K17-K52 [Stollar *et al.*, 2003]), tiene como consecuencia incrementos en T_m de entre 4 y 11 °C, así como una prolongación de la hélice III de unos 4 residuos (Fausti *et al.*, 2001).

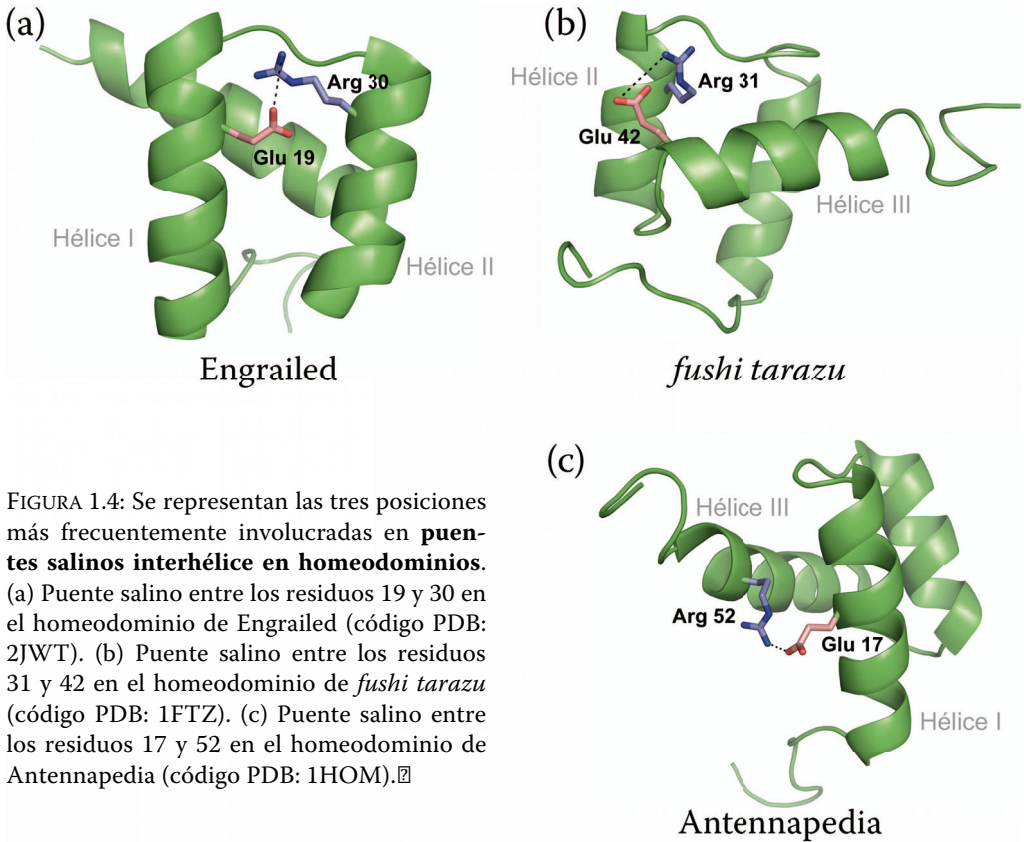


FIGURA 1.4: Se representan las tres posiciones más frecuentemente involucradas en **puentes salinos interhélice en homeodominios**. (a) Puente salino entre los residuos 19 y 30 en el homeodominio de Engrailed (código PDB: 2JWT). (b) Puente salino entre los residuos 31 y 42 en el homeodominio de *fushi tarazu* (código PDB: 1FTZ). (c) Puente salino entre los residuos 17 y 52 en el homeodominio de Antennapedia (código PDB: 1HOM).²

De un modo similar, los segmentos de hélice α I y II pueden estar conectados mediante un puente salino que involucra las posiciones 19-30 (Clarke, 1995; Isaac *et al.*, 1995; Stollar *et al.*, 2003) (Figura 1.4a). Dicha interacción está presente en un buen número de homeodominios, si bien presenta un grado de conservación inferior al de los contactos 31-42 y 17-52, anteriormente mencionados. La importancia del par 19-30 ha sido puesta de manifiesto mediante estudios de mutantes. Así, se ha demostrado que la sustitución E19K en *rough* conlleva una pérdida sustancial de función en *Drosophila* (Heberlein *et al.*, 1994). En ocasiones, residuos cargados en 19-30 pueden acompañarse de otros en las posiciones 15, 33 ó 37, posibilitando la formación de redes de puentes salinos entre las hélices I y II (Clarke *et al.*, 1994), tal como se representa en la Figura 1.5.

En lo referente a estos puentes salinos frecuentemente observados en homeodominios, es de destacar que en algunos casos involucran posiciones sólo parcialmente expuestas al disolvente, que empaquetan simultáneamente con el núcleo hidrófobo de la proteína (“posiciones fronterizas” según la

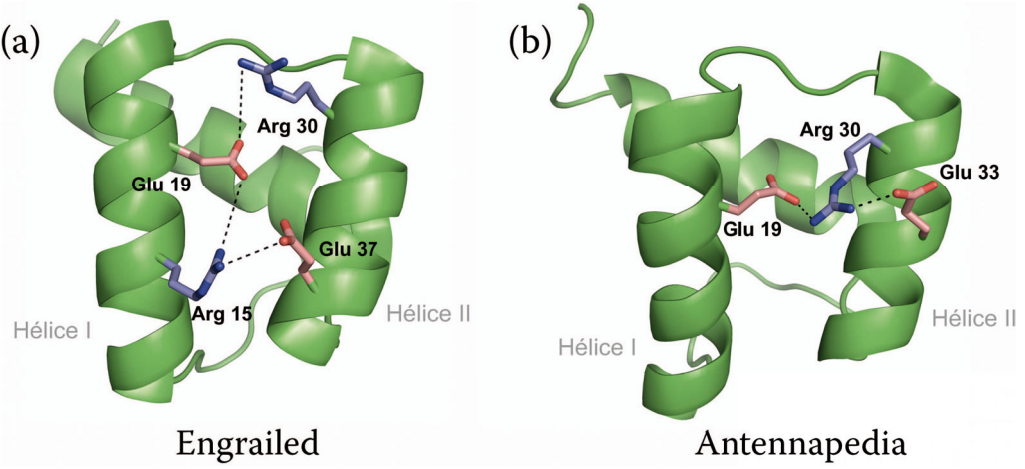


FIGURA 1.5: Se representan dos **redes alternativas de puentes salinos** que involucran el par 19-30, observadas en homeodominios naturales. (a) Red formada entre los residuos 19, 30, 15 y 37 en el homeodominio de Engrailed (código PDB: 2JWT). (b) Red formada entre las posiciones 19, 30 y 33 en el homeodominio de Antennapedia (código PDB: 1HOM).²

terminología de Mayo y colaboradores [Dahiyat y Mayo, 1997]). Éste es el caso de los residuos 15, 19, 30, 31 y 42. No resulta sorprendente comprobar que estas posiciones están, en un número reducido de secuencias, ocupadas por

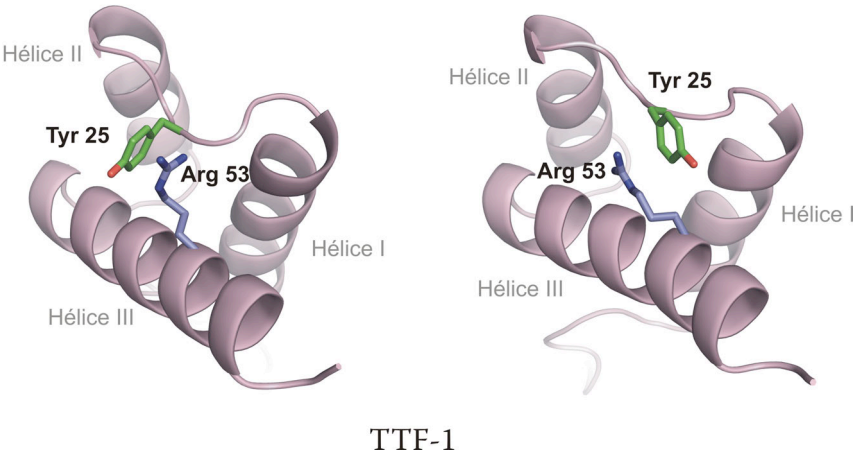


FIGURA 1.6: **Interacción catión- π** entre la **tirosina 25** y la **arginina 53**, observada en el homeodominio de TTF-1 (código PDB: 1FTT). De acuerdo con los datos de RMN, dicho contacto puede presentar diferentes geometrías.²

²

aminoácidos alifáticos. La abundancia de residuos cargados en dichas regiones (donde el coste energético asociado a la desolvatación parcial de la carga debe de ser sustancial) resulta llamativa y su papel en la estabilidad del homeodominio constituirá uno de los puntos tratados en esta tesis.

Finalmente, otro contacto muy conservado en homeodominios es la interacción catión- π entre el residuo Y25 del lazo I-II y la arginina 53 de la hélice III (Figura 1.6). Además de su aparente papel estabilizante, esta interacción podría posibilitar la correcta colocación de los dos residuos para contactar con el ADN. Ambas conclusiones se extraen, una vez más, de estudios con mutantes. Así pues, la mutación R53C en HESX-1 causa un pronunciado descenso de 10 °C en T_m e induce una pérdida casi completa de afinidad por ADN (de la Mata *et al.*, 2002). Por otro lado, los mutantes de *Drosophila* Y25N y el aparentemente conservativo R53K, en la homeoproteína *rough*, exhiben severos fenotipos de pérdida de función (Heberlein *et al.*, 1994).

1.1.3. Estabilidad de los homeodominios

Dada la gran relevancia biológica de los homeodominios, no resulta sorprendente que su estabilidad, así como los factores que la determinan, hayan sido objeto de un buen número de estudios en los últimos veinte años. Éstos comprenden sistemas de orígenes diversos tales como levadura, mosca, pequeños mamíferos y, finalmente, humanos. Las conclusiones fundamentales que pueden deducirse de un análisis comparado de los resultados se resumen en el presente apartado.

Como se puede comprobar en la Tabla 1.1, todos los homeodominios descritos hasta la fecha presentan valores de T_m de entre 25 y 60 °C. Dichas temperaturas son bajas si las comparamos con las descritas para otros sistemas, incluidas proteínas de tamaño similar así como otros motivos de reconocimiento de ADN (Frankel *et al.*, 1987; Fields, 2001; Di Nardo *et al.*, 2003; Kumar *et al.*, 2003; Sakai-Kato *et al.*, 2009). Una conclusión que se deriva de estos datos es que, en un buen número de casos, el homeodominio tiene una estabilidad marginal a las temperaturas a las que funciona *in vivo*. A modo de ejemplo, en *Drosophila*, los homeodominios de *fushi tarazu* y *vnd* (NK-2) presentan valores de T_m (32 y ~34 °C, respectivamente [Qian *et al.*, 1994a; Tsao *et al.*, 1994; Weiler *et al.*, 1998; Gonzalez *et al.*, 2001; Dragan *et al.*, 2006]) próximos a la temperatura de su entorno (20-25 °C). Más destacables aún son los casos de TTF-1, Oct-2 POU e ISL-1 humanos, cuyas temperaturas de desnaturalización son ~40, ~35 y 31 °C, respectivamente (Damante *et al.*, 1994; Sivaraja *et al.*, 1994; Behravan *et al.*, 1997; Tell *et al.*, 1999; Del Vecchio *et al.*, 2008). Resulta sorprendente comprobar que dichos valores son similares o

TABLA 1.1: Datos termodinámicos de estabilidad de algunos homeodominios representativos^a

Homeodominio	T_m (°C)	ΔG_F (kcal/mol) ^b	ΔC_p (kcal/[mol·K])	Tampón ^c	Referencias
Antennapedia	48	-2,7 (24 °C) ^{d, e}	-	A, B, C	Otting <i>et al.</i> , 1988; Qian <i>et al.</i> , 1994a; Tell <i>et al.</i> , 1999
	60	-	-	D	Güntert <i>et al.</i> , 1991
	55,5 ^f	-5,25 (20 °C) ^f	-0,29 ^f	E	Dragan <i>et al.</i> , 2006
Bicoid	44	-	-	F	Subramaniam <i>et al.</i> , 2001
Engrailed	55,5	-	-	G	Ades y Sauer, 1994
	46	-	-	H	Clarke <i>et al.</i> , 1994
	52,1 ^f	-2,2 (20 °C) ^e	-0,7	I	Stollar <i>et al.</i> , 2003
	48,7 ^f	-2,79 (20 °C) ^f	-0,36 ^f	E	Dragan <i>et al.</i> , 2006
<i>fushi tarazu</i>	27	-	-	B	Qian <i>et al.</i> , 1994a
ISL-1	31	-	-	J	Behravan <i>et al.</i> , 1997
	30	-	-	K	Ippel <i>et al.</i> , 1998
MATα2	56,4 ^f	-	-0,24 ^f	L	Carra y Privalov, 1997
	52,2 ^f	-3,92 (20 °C) ^f	-0,45 ^f	E	Dragan <i>et al.</i> , 2006
NKX2.5	53,9	-2,4 (25 °C) ^e	-0,14 ^f , -1,2	M	Fodor <i>et al.</i> , 2005
TTF-1	39	-1,5 (24 °C) ^e	-	C	Tell <i>et al.</i> , 1999
	42	-	-0,36 ^f , -0,38	N	Del Vecchio <i>et al.</i> , 2008
vnd (NK-2)	25	-	-	Ñ, O	Tsao <i>et al.</i> , 1994; Weiler <i>et al.</i> , 1998
	43,0	-1,2 (25 °C)	-0,52	P	Gonzalez <i>et al.</i> , 2001
	42,1 ^f	-2,25 (20 °C) ^f	-0,38 ^f	E	Dragan <i>et al.</i> , 2006

^a Salvo que se indique con los superíndices *e* o *f*, los valores fueron derivados de experimentos de desnaturalización térmica por dicroísmo circular.

^b Entre paréntesis se indica la temperatura a la que se midió ΔG_F .

^c Los tampones usados son los siguientes: A, no indicado, pH 4,5; B, no indicado, pH 4,8; C, 10 mM fosfato pH 7,5, 5 mM NaF; D, no indicado, pH 6,9; E, 20 mM acetato pH 5,0, 100 mM NaCl; F, 10 mM fosfato pH 7,0, 50 mM NaCl; G, 50 mM fosfato pH 7,0, 100 mM KCl; H, 10 mM tris-HCl pH 7,4; I, 50 mM acetato pH 5,7, 100 mM NaCl; J, 20 mM fosfato pH 6,0, 50 mM NaCl; K, 25 mM fosfato pH 6,0, 25 mM NaCl; L, 20 mM glicina pH 5,0, 100 mM NaCl; M, 20 mM MOPS pH 7,4, 100 mM NaCl; N, 20 mM acetato pH 5,0, 20 mM KCl; Ñ, no indicado; O, 20 mM acetato pH 4,5; P, 50 mM HEPES pH 7,4, 100 mM NaCl. Cuando se dan varios tampones en una misma fila, se indican en el mismo orden que las referencias correspondientes.

^d Valor obtenido de Tell *et al.*, 1999, con el tampón C.

^e Valores derivados de experimentos de desnaturalización química con urea por dicroísmo circular.

^f Valores derivados de experimentos de desnaturalización térmica por calorimetría diferencial de barrido (DSC).

incluso inferiores a la temperatura corporal del hombre, lo que sugiere que, en estas condiciones, los homeodominios estarían en una buena proporción desnaturalizados.

En concordancia con la baja estabilidad térmica de los homeodominios, sus valores de energía libre de plegamiento (ΔG_F) se sitúan típicamente entre -1 y -5 kcal/mol a 20-25 °C, lo cual es equiparable a la energía de un único

punto de hidrógeno. Estos valores son el resultado de un balance entre términos entálpicos (ΔH_F) y entrópicos (ΔS_F). Como sucede en todas las proteínas, se ha observado que ΔH_F en homeodominios favorece el estado plegado y ΔS_F se opone a él (Fodor *et al.*, 2005). Mayores particularidades se encuentran al analizar el ΔC_p . Así pues, pese a que se han descrito valores de entre $-0,1$ y $-1,4$ kcal/(mol·K), la mayoría de las mediciones arrojan valores en torno a $-0,4$ kcal/(mol·K), esto es, más bajos de lo esperado en base al tamaño del homeodominio (unos $-0,8$ kcal/[mol·K] para proteínas de 60 residuos [Myers *et al.*, 1995]). El origen de esta tendencia reside en buena medida en que la capacidad calorífica absoluta del estado nativo es mayor de lo normal, lo cual puede atribuirse al desorden intrínseco de las regiones N- y C-terminal de estas proteínas, así como a la escasa rigidez de su plegamiento.

Lejos de representar un hecho meramente anecdótico, o incluso una desventaja evolutiva, la estabilidad marginal de este importante motivo de reconocimiento de ADN podría presentar implicaciones funcionales de calado. Así por ejemplo, en los casos de Oct-2 POU e ISL-1 anteriormente mencionados, se ha especulado que la proteína sólo adquiere su plegamiento funcional como consecuencia de su interacción con el ADN diana (Ippel *et al.*, 1998). En este caso, el proceso de reconocimiento constituiría un ejemplo extremo de ajuste inducido (*induced fit*) (tal como se discutirá más adelante, este modelo es aplicable, si bien en menor medida, a todos los homeodominios). Las consecuencias termodinámicas de este modo de unión han sido objeto de numerosos análisis y se describirán con más detalle en el apartado 1.2.4.

Por otro lado, parece lógico suponer que la baja estabilidad de los homeodominios determina un tiempo de vida corto de los mismos en el medio intracelular. Se ha especulado que en proteínas reguladoras ésta sería una propiedad esencial, permitiendo que la célula respondiese de manera rápida y eficaz a los cambios constantes en el entorno (Cañadillas *et al.*, 2006). De acuerdo con esta idea, ciertas proteínas reguladoras habrían evolucionado para ser relativamente inestables en condiciones fisiológicas.

Diversos estudios han puesto de manifiesto la estrecha relación existente entre la estabilidad de los homeodominios y su competencia funcional *in vivo*. Por ejemplo, ciertos mutantes en el residuo 52, que conllevan la formación o rotura del puente salino estabilizante E17-R52 (Figura 1.4c), suelen mostrar un comportamiento peculiar en *Drosophila*. Así pues, se ha descrito que la estabilización lograda por la mutación H52R en vnd (NK-2) resulta tóxica para la mosca (Fodor *et al.*, 2005). De un modo similar, la desestabilización promovida por la sustitución R52H en *even-skipped* (Frasch *et al.*, 1988), o R52L y R52C en *rough* (Heberlein *et al.*, 1994), son igualmente nocivas para el organismo dando lugar a un curioso comportamiento: el desarrollo del mismo se torna dependiente de la temperatura de crecimiento (a 16-18 °C el fenotipo

es normal, y al ir subiendo hasta 29 °C aparecen y se incrementan diversas malformaciones). Estas observaciones experimentales sugieren que la estabilidad debe mantenerse dentro de unos márgenes muy definidos, por lo cual dicho parámetro habría estado finamente ajustado por la evolución.

CAUSAS GLOBALES DE LA BAJA ESTABILIDAD DE LOS HOMEODOMINIOS

En los anteriores párrafos hemos apuntado las posibles implicaciones funcionales que se derivan de la baja estabilidad de los homeodominios. En este subapartado, se resumen brevemente las razones estructurales que determinan esta propiedad.

En primer lugar, tal como se apuntó en el apartado 1.1.1, de los 60 residuos que integran este motivo estructural, aproximadamente el 25% está total o parcialmente desestructurado en disolución (brazo N-terminal y hélice III-b) (Figura 1.7). De acuerdo con esto, la región plegada comprende alrededor de 45 aminoácidos y determina un núcleo hidrófobo de reducidas dimensiones,

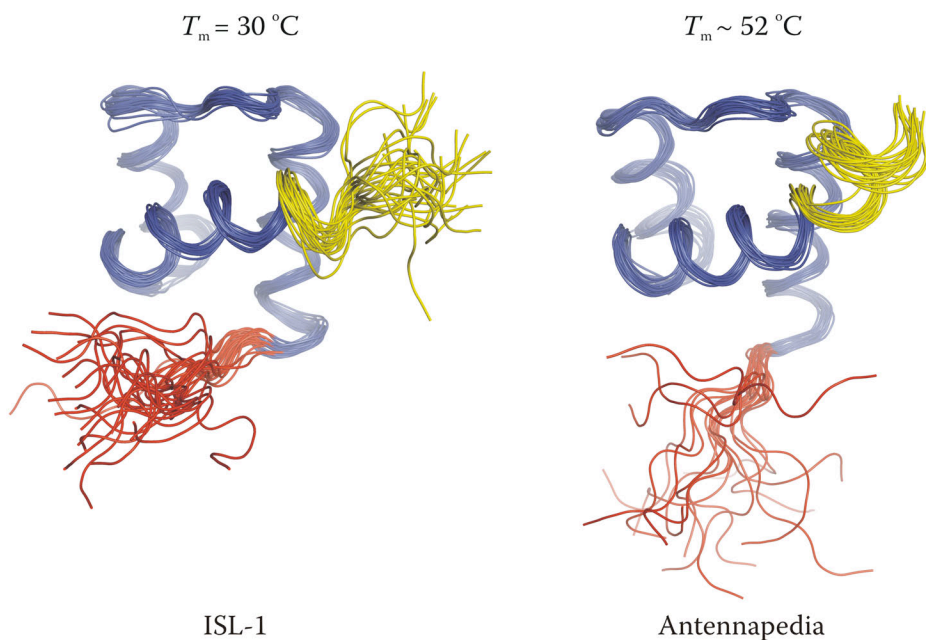


FIGURA 1.7: Regiones desordenadas de los homeodominios, ilustradas mediante la superposición de 25 y 19 estructuras de RMN de ISL-1 (código PDB: 1BW5) y Antennapedia (código PDB: 1HOM), respectivamente. En rojo se representa el brazo N-terminal, y en amarillo la hélice III-b. Se indican las T_m de ambas proteínas. Puede apreciarse como, cualitativamente, dicho parámetro guarda proporción con el grado de estructuración de la hélice III-b.

cuyo carácter parcialmente dinámico ha sido puesto de manifiesto por diversos estudios (Damante *et al.*, 1994; Qian *et al.*, 1994a; Nanda y Brand, 2000; Fodor *et al.*, 2005). Estas características estructurales constituyen la primera de las dos causas más evidentes que determinan una baja estabilidad en homeodominios.

A pesar de que las regiones desestructuradas se localizan en todos los casos en zonas idénticas, existen sutiles diferencias entre homeodominios, que afectan de modo más significativo a la extensión de la hélice III-b. Resulta interesante comprobar que este parámetro correlaciona en buena medida con la estabilidad del plegamiento (Figura 1.7). Así pues, Engrailed y sobre todo Antennapedia, que presentan una hélice III-b bastante ordenada, son de los homeodominios más estables, con T_m s en torno a los 50 °C (Tell *et al.*, 1999; Mayor *et al.*, 2000; Stollar *et al.*, 2003; Dragan *et al.*, 2006). En cambio, casos como los de vnd (NK-2), o ISL-1, con una hélice III-b desestructurada (Phillips *et al.*, 1991; Ippel *et al.*, 1998), presentan, como señalábamos anteriormente, estabilidades muy bajas, que pueden incrementarse sustancialmente mediante la introducción de mutaciones adecuadas. De este modo, la doble sustitución H52R/T56W en vnd (NK-2) conlleva una estructuración de la hélice III-b, acompañada de un incremento en T_m de 8 °C (Weiler *et al.*, 1998; Gonzalez *et al.*, 2001). En el apartado 1.1.2 hemos señalado las interacciones estabilizantes que pueden establecer los residuos en posiciones 52 y 56, que originan prolongaciones de la hélice III de este tipo.

Además de no contribuir a la estabilidad del plegamiento, se ha demostrado que, sorprendentemente, las regiones desorganizadas del polipéptido, tales como el segmento N-terminal, constituyen en sí mismas un factor de desestabilización. Así, se ha observado una mejora de ΔG_F de 0,5 kcal/mol a 20 °C al eliminar el brazo N-terminal de Antennapedia y de vnd (NK-2), equivalente a un aumento en T_m de 3,5 °C (Dragan *et al.*, 2006). Este efecto, cuyo origen se desconoce, es insensible a la fuerza iónica del entorno y, por lo tanto, no parece atribuible a interacciones electrostáticas de largo alcance.

La segunda razón que justifica la baja estabilidad relativa de los homeodominios surge de restricciones de tipo funcional y está relacionada con su elevada carga neta positiva, situada típicamente entre +4 y +13 a pH fisiológico. Como es lógico, esta propiedad tiene su origen en la presencia de un gran número de residuos básicos, esenciales para el reconocimiento de ADN, muchos de los cuales se acumulan en puntos concretos de la estructura (Stollar *et al.*, 2003). Evidentemente, la repulsión electrostática derivada de esta concentración de carga es superior en el estado nativo, al ser más compacto que el desplegado, lo que explica su efecto desestabilizante. De acuerdo con esto, resulta interesante comprobar como diversas versiones termoestables del homeodominio de Engrailed descritas en la literatura, resultantes de la introducción de múltiples mutaciones, presentan un rasgo común poco

frecuente en versiones naturales de estas proteínas: una carga neta cercana a cero (Marshall *et al.*, 2002; Zollars *et al.*, 2006; Shah *et al.*, 2007).

Finalmente, desde un punto de vista estructural, parece lógico suponer que el patrón de interacciones terciarias presente en los homeodominios ha sido optimizado por la evolución para limitar la estabilidad de estas proteínas, manteniendo este parámetro en los márgenes adecuados para el desempeño de su función biológica.

MÉTODOS PARA AUMENTAR LA ESTABILIDAD DE LOS HOMEODOMINIOS

De acuerdo con lo apuntado en el subapartado anterior, una de las estrategias más evidentes para la estabilización de homeodominios se basaría en la optimización de su balance electrostático, lo cual puede conseguirse, de un modo sencillo, modificando ciertas condiciones de su entorno, como son el pH y la fuerza iónica.

Por lo que respecta al efecto del pH, el carácter policationico del homeodominio determina, típicamente, valores altos de punto isoeléctrico (entre 10,0 y 11,7). En consecuencia, elevar el pH disminuye la carga neta y, por norma general, estabiliza estas proteínas. Como ejemplos, en la [Figura 1.8a](#) podemos ver el caso del homeodominio de *fushi tarazu*, cuya T_m asciende 9 °C al pasar de pH 3,5 a pH 6,1 (Qian *et al.*, 1994a), o el de MAT α 2, que experimenta una subida de 10 °C en T_m al elevar el pH de 3,5 a 7,0 (Carra y Privalov, 1997). En general, esta dependencia del pH suele ser especialmente acusada por debajo de pH 6,0 siendo menor o incluso inexistente en medios neutros (Carra y Privalov, 1997; Fodor *et al.*, 2005; Del Vecchio *et al.*, 2008). Finalmente, es preciso mencionar que, a pesar de no constituir sus condiciones óptimas de estabilidad, diversos estudios estructurales y termodinámicos de homeodominios se han realizado a pH ácido, con objeto de favorecer su solubilidad (Otting *et al.*, 1988; Qian *et al.*, 1994a; Dragan *et al.*, 2006).

En cuanto a la fuerza iónica, su efecto estabilizante sobre estas proteínas ha sido puesto de manifiesto por diversos estudios ([Figura 1.8b](#)). Así por ejemplo, a pHs parecidos, la T_m del homeodominio de MAT α 2 de *Saccharomyces* aumenta 8 °C al añadir 400 mM NaCl (Carra y Privalov, 1997); en el caso de Antennapedia, la T_m sube 20 °C al añadir 500 mM NaCl (Dragan *et al.*, 2006); por último, para vnd (NK-2), la T_m experimenta un marcado ascenso de 26 °C al añadir 500 mM NaCl (Dragan *et al.*, 2006). Es de destacar que en estos tres ejemplos existe una correlación entre la magnitud de la estabilización promovida por la fuerza iónica y la carga neta del polipéptido (+9, +11 y +13, respectivamente). También se ha comprobado como, claramente, la estabilización del homeodominio por fuerza iónica es más

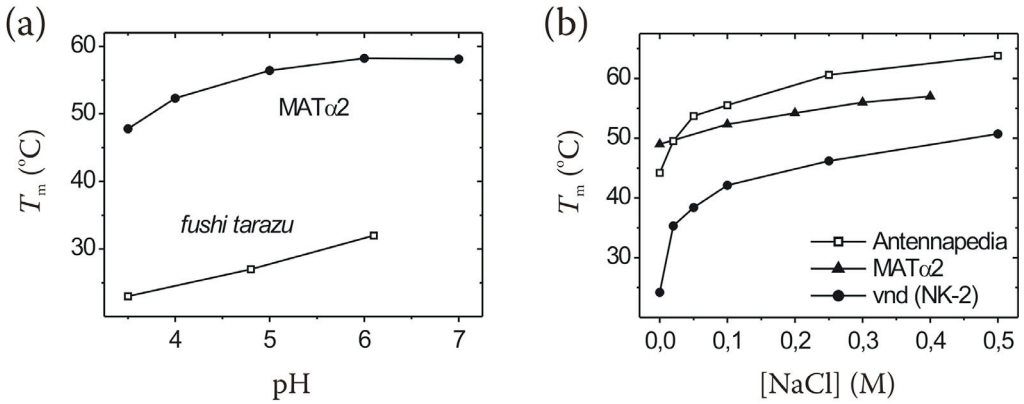


FIGURA 1.8: **Estabilización de los homeodominios** mediante el aumento del pH o de la fuerza iónica. (a) Dependencia de T_m con el pH para los homeodominios de MAT α 2 (Carra y Privalov, 1997) y *fushi tarazu* (Qian *et al.*, 1994a). (b) Dependencia de T_m con [NaCl] para los homeodominios de Antennapedia (Dragan *et al.*, 2006), MAT α 2 (Carra y Privalov, 1997) y vnd (NK-2) (Dragan *et al.*, 2006).

acusada a pHs bajos (Carra y Privalov, 1997; Gonzalez *et al.*, 2001; Dragan *et al.*, 2006).

Empleando una aproximación mutacional es posible obtener efectos similares o superiores sobre el ΔG_F de estos sistemas. De hecho, el carácter marginalmente estable de los homeodominios (cuya posible relevancia funcional ha sido anteriormente apuntada) permite intuir la existencia de un amplio margen de mejora mediante las sustituciones adecuadas.

En concordancia con estas observaciones, diversos trabajos describen apreciables aumentos de estabilidad conseguidos mediante un número reducido de sustituciones (entre una y tres). Algunos de estos ejemplos, tales como la sustitución H52R/T56W en vnd (NK-2), o K52E en Engrailed, han sido recogidos en apartados anteriores (Weiler *et al.*, 1998; Gonzalez *et al.*, 2001; Stollar *et al.*, 2003). Otros ejemplos incluyen las mutaciones Y54M en TTF-1, que implica un incremento de T_m de 3 °C (Tell *et al.*, 1999), Q50A o K52A en Engrailed, con aumentos de T_m de 1 y 10 °C, respectivamente (Ades y Sauer, 1994; Stollar *et al.*, 2003), o F8Y en Bicoid, que supone un incremento en T_m de 2 °C (Subramaniam *et al.*, 2001).

Evidentemente, resultados más acusados pueden obtenerse mediante un completo rediseño del patrón de interacciones presentes en la proteína. Utilizando esta aproximación, el grupo del profesor Mayo ha obtenido variantes termoestables del homeodominio de Engrailed (T_m de la proteína salvaje: 49 °C). Así, una optimización de las interacciones electrostáticas, restringida a un grupo de posiciones clave, condujo a un homeodominio con 23

sustituciones y una T_m de 88 °C (Marshall *et al.*, 2002). Una nueva optimización posterior de la totalidad de los residuos dio lugar a dos variantes que presentaban 39 y 40 mutaciones y una T_m superior a 99 °C (Shah *et al.*, 2007). Finalmente, un posicionamiento adecuado de residuos hidrófobos y polares en posiciones enterradas y expuestas al disolvente, respectivamente, dio lugar a una variante de Engrailed hipertermoestable, con una T_m de 114 °C (Marshall y Mayo, 2001). En este último trabajo se dedicó especial atención a las posiciones fronterizas, ejemplos de las cuales hemos visto en el apartado 1.1.2. Aumentar el número de residuos apolares en dichas posiciones daba lugar a variantes de mayor estabilidad, si bien tenía como contrapartida una disminución en la solubilidad de la proteína y en su especificidad conformacional.

Es importante señalar que en estos estudios se consiguió un incremento espectacular de estabilidad sin tener en consideración la funcionalidad del sistema. De hecho, un buen número de las sustituciones empleadas afectan a posiciones clave en el reconocimiento de ADN, que evidentemente no han sido optimizadas por la evolución atendiendo a un criterio de estabilidad.

1.2

UNIÓN DEL HOMEODOMINIO A ADN

En general, los factores de transcripción desempeñan su función biológica mediante su unión, altamente específica, a la secuencia correcta de ADN. Naturalmente, dicha especificidad tiene su origen en unas interacciones concretas entre la proteína y el ácido nucleico. En los siguientes apartados se abordan brevemente los rasgos generales de la asociación específica homeodominio-ADN, así como las propiedades y relevancia funcional de los complejos no específicos.

1.2.1. Secuencia de reconocimiento específico de homeodominios

Por norma general, la secuencia diana de los homeodominios consta de un núcleo de cuatro bases. Para la mayoría, incluido HESX-1, dicho núcleo consiste en la tétrada 5'-TAAT-3', si bien ésta también puede ser 5'-CAAG-3' u otras (Gehring *et al.*, 1994b; Weiler *et al.*, 1998; Dave *et al.*, 2000). A estas cuatro primeras bases, esenciales para la unión de alta afinidad, les siguen generalmente dos que modulan la especificidad de secuencia y presentan una mayor variabilidad en secuencias de reconocimiento de estas proteínas (Laughon, 1991; Ekker *et al.*, 1992; Catron *et al.*, 1993; Wilson *et al.*, 1996) (Tabla 1.2).

Es de destacar la existencia de homeodominios capaces de reconocer no sólo una, sino varias secuencias de forma específica, como es el caso de Bicoid (Dave *et al.*, 2000) o de TTF-1 (Fabbro *et al.*, 1996). Esto es posible gracias a la elevada adaptabilidad estructural del polipéptido, especialmente en el entorno de las posiciones 50 y 54, que son clave para la especificidad (Wilson *et al.*, 1996; Dave *et al.*, 2000).

TABLA 1.2. Secuencias de ADN de unión específica de algunos homeodominios representativos

Organismo	Homeodominio	Secuencias de reconocimiento (5'→3') ^a	Referencias
<i>Saccharomyces</i>	MATα2	TTACAT	Wolberger <i>et al.</i> , 1991; Gruschus <i>et al.</i> , 1997
<i>Drosophila</i>	Antennapedia	TAATGG	Laughon, 1991; Billeter <i>et al.</i> , 1993; Fraenkel y Pabo, 1998
	Bicoid	TAATCC TAAGCT	Laughon, 1991; Wilson <i>et al.</i> , 1996 Dave <i>et al.</i> , 2000
	Engrailed	TAATTA	Kissinger <i>et al.</i> , 1990; Laughon, 1991; Ades y Sauer, 1994; Sato <i>et al.</i> , 2004
	<i>even-skipped</i>	TAATTG	Laughon, 1991; Hirsch y Aggarwal, 1995
	<i>fushi tarazu</i>	TAATTG	Laughon, 1991; Gehring <i>et al.</i> , 1994b
	Paired	TAATCG	Treisman <i>et al.</i> , 1989; Laughon, 1991
	vnd (NK-2)	CAAGTG	Gruschus <i>et al.</i> , 1999; Gonzalez <i>et al.</i> , 2001
Humano	HESX-1	TAATTG	Brickman <i>et al.</i> , 2001; de la Mata <i>et al.</i> , 2002
	ISL-1	TAATAT	Ippel <i>et al.</i> , 1998
	MSX-1	TAATTG	Isaac <i>et al.</i> , 1995; Hovde <i>et al.</i> , 2001
	PITX2 (RIEG1)	TAATCC	Amendt <i>et al.</i> , 1999; Banerjee-Basu <i>et al.</i> , 2003
	TTF-1	CAAGTG	Esposito <i>et al.</i> , 1996; Fabbro <i>et al.</i> , 1996; Del Vecchio <i>et al.</i> , 2008
		GGCTCA	Fabbro <i>et al.</i> , 1996

^a En negrita se señala el núcleo de cuatro bases de la secuencia. Para los homeodominios de Bicoid y TTF-1, se muestran en gris secuencias alternativas de unión de alta afinidad.

1.2.2. Características estructurales de los complejos específicos homeodominio/ADN

En la [Figura 1.9](#) se representa una vista general de la estructura tridimensional del homeodominio de Engrailed asociado a un oligonucleótido de ADN con su secuencia diana. En este apartado vamos a detallar las características estructurales de estos complejos específicos.

Por lo que respecta al ADN, éste mantiene su estructura tipo B en el complejo, si bien experimenta en la mayoría de los casos una ligera curvatura por efecto de la asociación (Wolberger, 1996; Gruschus *et al.*, 1999; Ke *et al.*, 2002), así como muy ligeras desviaciones de los parámetros geométricos típicos de la forma B (Kissinger *et al.*, 1990; Wolberger *et al.*, 1991; Gruschus *et al.*, 1999). Además, en algunos casos los surcos mayor y menor se ensanchan o

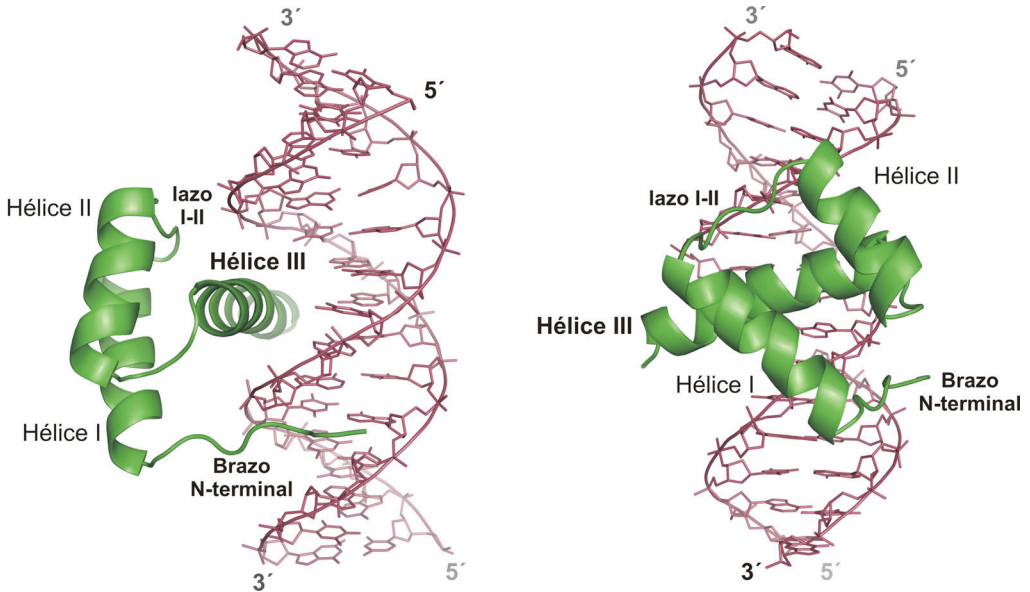


FIGURA 1.9: **Estructura tridimensional del homeodominio de Engrailed unido al ADN**, ilustrada en dos perspectivas diferentes (código PDB: 1HDD). En negrita se etiquetan las zonas del homeodominio más importantes en el reconocimiento. El ADN se representa en rosa.

estrechan para adaptarse mejor a las regiones de reconocimiento del homeodominio (Kissinger *et al.*, 1990; Li *et al.*, 1995; Gruschus *et al.*, 1997, 1999; Grant *et al.*, 2000).

Por lo que respecta a la proteína, numerosos estudios han demostrado que, a pesar de que la topología de plegamiento permanece esencialmente inalterada en el estado asociado (Figura 1.9), la formación del complejo va acompañada de ajustes estructurales significativos en dos regiones concretas del homeodominio (Figura 1.10):

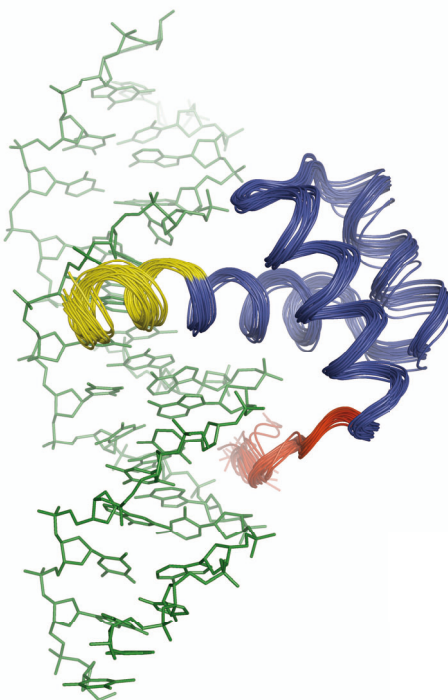
- La hélice III, también llamada hélice de reconocimiento, se inserta en el surco mayor del ADN, lo cual implica la estructuración de la porción III-b que, como hemos visto, en el estado libre se encuentra parcialmente desordenada (Morita *et al.*, 1993; Damante *et al.*, 1994; Tsao *et al.*, 1994; Carra y Privalov, 1997; Fausti *et al.*, 2001; Fodor *et al.*, 2005; Del Vecchio *et al.*, 2008). De este modo, la hélice III se muestra más larga en el estado asociado, y su división en III-a y III-b no es en este caso aparente.
- El brazo N-terminal, flexible y completamente desestructurado en el estado libre, adquiere una estructura definida en el complejo, estableciendo

importantes contactos con el surco menor del ADN (Kissinger *et al.*, 1990; Billeter *et al.*, 1993; Hirsch y Aggarwal, 1995; Gruschus *et al.*, 1997).

Además de estas dos regiones, el núcleo hidrófobo y otras zonas del homeodominio podrían sufrir sutiles transiciones conformacionales durante el reconocimiento específico, permitiéndole amoldarse a la superficie del ADN.

Tomado de un modo global, estas observaciones son consistentes con un modelo de reconocimiento por ajuste inducido (*induced fit*).

FIGURA 1.10: **Estructuración del brazo N-terminal y de la hélice III-b como consecuencia de la unión a ADN**, ilustrada mediante la superposición de 20 estructuras de RMN del homeodominio de vnd (NK-2) asociado a ADN (código PDB: 1NK2). En rojo se representa el brazo N-terminal, y en amarillo la hélice III-b. El ADN se muestra en verde. [?]



1.2.3. Interacciones observadas en los complejos específicos homeodominio/ADN

De acuerdo con los datos disponibles, tres porciones concretas del homeodominio juegan un papel clave en el reconocimiento de ADN y en la estabilización del complejo:

a) **Hélice III:** Esta región del polipéptido participa en diversos contactos con las dos hebras de la doble hélice, tanto directos como mediados por moléculas de agua que se localizan en la interfase del complejo. Entre ellos, cabe destacar

los establecidos por residuos catiónicos altamente conservados, situados en las posiciones 46, 52, 53, 55 y 57 (Kissinger *et al.*, 1990; Wolberger *et al.*, 1991; Billeter *et al.*, 1993; Hirsch y Aggarwal, 1995; Wilson *et al.*, 1996; Gruschus *et al.*, 1997; Stollar *et al.*, 2003). Dichas posiciones contribuyen a la estabilidad del complejo mediante interacciones polares y/o electrostáticas, de corto o medio alcance, con los fosfatos del esqueleto del ADN, y en muchos casos mediante puentes de hidrógeno (Figura 1.11b).

Adicionalmente, se observan otras interacciones proteína-ADN clave, tanto polares como de naturaleza hidrófoba, que involucran las posiciones 47, 50, 51 y 54 del homeodominio (Figura 1.11c) (Kissinger *et al.*, 1990; Otting *et al.*, 1990; Wolberger *et al.*, 1991; Billeter *et al.*, 1993; Li *et al.*, 1995; Gruschus *et al.*, 1997, 1999; Fraenkel y Pabo, 1998). Estos residuos establecen numerosos contactos con las unidades de desoxirribosa, así como con las bases nitrogenadas. Por ello, se les considera en buena medida responsables de la especificidad de secuencia del polipéptido. Teniendo esto en cuenta, no resulta sorprendente comprobar que las posiciones 47, 50 y 54 presentan, en general, una alta variabilidad al pasar de un homeodominio a otro (Treisman *et al.*, 1989; Laughon, 1991; Ades y Sauer, 1994; Wilson *et al.*, 1996; Weiler *et al.*, 1998; Tell *et al.*, 1999). Entre ellas, destaca por su importancia el residuo 50, que juega un papel clave en la determinación fina de especificidad de secuencia. Esto lo consigue por medio de contactos con las dos bases de la secuencia diana que siguen al núcleo 5'-TAAT-3'. Como se ha mencionado anteriormente, dicha región del ADN es la que presenta mayor variabilidad en secuencias de reconocimiento de los distintos homeodominios (Treisman *et al.*, 1989; Kissinger *et al.*, 1990; Otting *et al.*, 1990; Wilson *et al.*, 1996; Gruschus *et al.*, 1999).

A diferencia de los residuos 47, 50 y 54, la posición 51 incorpora una asparagina altamente conservada, que contacta con el propio núcleo de cuatro bases de la secuencia de reconocimiento, invariable en la mayoría de los casos (Kissinger *et al.*, 1990; Hirsch y Aggarwal, 1995; Gruschus *et al.*, 1997, 1999; Fraenkel y Pabo, 1998; Grant *et al.*, 2000).

Curiosamente, diversos estudios han puesto de manifiesto que algunos residuos de la hélice III son relevantes para el reconocimiento de ADN, a pesar de no contactar con él directamente. Tal es el caso de la posición 56, que podría tener un papel en definir la especificidad de secuencia (Qian *et al.*, 1994a; Weiler *et al.*, 1998), así como el triptófano conservado en posición 48, que parece imprescindible para mantener las relaciones espaciales del núcleo hidrófobo necesarias para el correcto posicionamiento de grupos clave en la formación del complejo (Kissinger *et al.*, 1990; Subramaniam *et al.*, 2001). Este triptófano, además, contacta con un fosfato del ADN a través de unas moléculas de agua observadas en varias estructuras cristalográficas (Figura

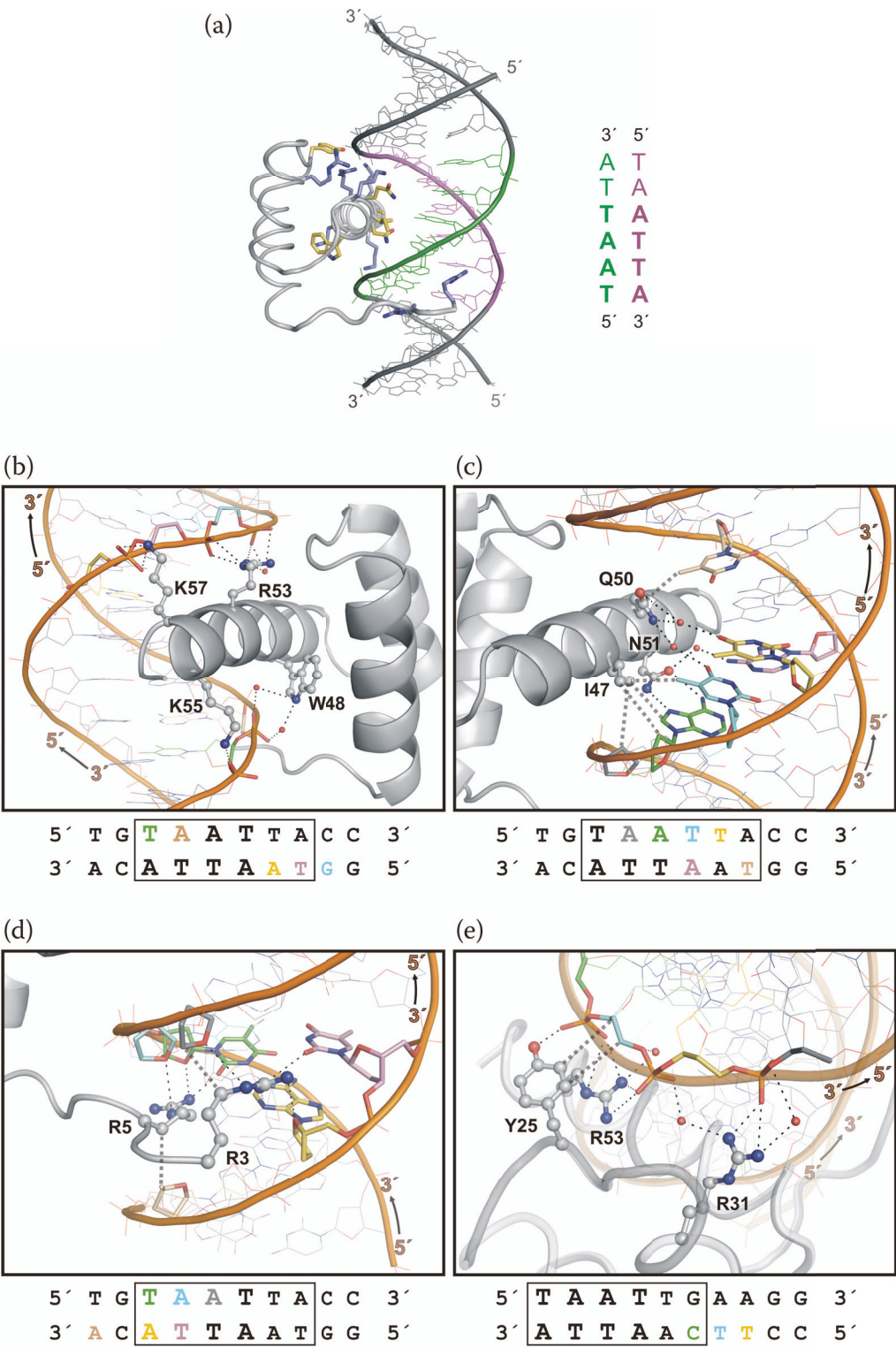


FIGURA 1.11 (PÁGINA ANTERIOR): **Interacciones intermoleculares más relevantes observadas en complejos homeodominio/ADN.** (a) Visión general del homeodominio de Engrailed unido al ADN (código PDB: 1HDD), en la que se destacan los residuos más importantes de los mencionados en el texto principal: en azul, de izquierda a derecha, R31, K52, R53, K46, K57, K55, R5 y R3; en amarillo, de arriba a abajo, Y25, Q50, A54, I47, N51 y W48. El ADN se muestra en negro, destacando en color la región comprendida por la secuencia de reconocimiento de Engrailed. Ésta se muestra en paralelo a la derecha, con el núcleo TAAT en negrita. (b) Complejo Engrailed/ADN (código PDB: 3HDD): detalle de las interacciones establecidas por los residuos de la hélice III W48, R53, K55 y K57 con los fosfatos de la doble hélice. (c) Complejo Engrailed/ADN (código PDB: 3HDD): detalle de los contactos establecidos por los residuos de la hélice III I47, Q50 y N51 con los azúcares y las bases de la doble hélice. (d) Complejo Engrailed/ADN (código PDB: 1HDD): detalle de los contactos establecidos por los residuos del brazo N-terminal R3 y R5 con los azúcares y las bases de la doble hélice. (e) Complejo MSX-1/ADN (código PDB: 1IG7): detalle de los contactos establecidos por los residuos Y25, perteneciente al lazo I-II, y R31, perteneciente a la hélice II, con el esqueleto azúcar-fosfato del ADN. Adicionalmente, se representa R53 con objeto de ilustrar su relación con Y25. En (b), (c), (d) y (e), las interacciones polares se indican con líneas negras discontinuas, y los contactos hidrófobos y de van der Waals con líneas grises discontinuas más anchas. Las moléculas de agua que median ciertos contactos polares se representan como puntos rojos. Debajo de cada panel, se indica la secuencia del ADN correspondiente, con la diana de reconocimiento encuadrada, y el núcleo TAAT destacado en mayor tamaño. Asimismo, los nucleótidos contactados se representan en el mismo color en que aparecen en la estructura. En todos los paneles, los átomos de oxígeno se representan en rojo, los de nitrógeno en azul, y los de fósforo en naranja.

1.11b) (Kissinger *et al.*, 1990; Hirsch y Aggarwal, 1995; Fraenkel y Pabo, 1998; Stollar *et al.*, 2003).

b) Brazo N-terminal: En segundo lugar, son de gran importancia los contactos establecidos por el brazo N-terminal con el surco menor del ADN, por conferir tanto afinidad como especificidad al proceso de unión. Éstos están mediados, en buena medida, por las posiciones 3 y 5 del polipéptido, que incorporan argininas en la mayoría de los casos (Gehring *et al.*, 1994a; Wolberger, 1996), si bien en posición 3 no es infrecuente encontrar lisina. De acuerdo con los datos estructurales disponibles, ambos residuos se insertan en el surco menor del dúplex de ADN, a la altura del núcleo de cuatro bases de la secuencia diana. Se establecen de este modo interacciones de puentes de hidrógeno con las bases nitrogenadas (Figura 1.11d) (Kissinger *et al.*, 1990; Billeter *et al.*, 1993; Hirsch y Aggarwal, 1995; Gruschus *et al.*, 1997, 1999; Fraenkel y Pabo, 1998), que contribuyen en buena medida a la especificidad de secuencia del homeodominio. Adicionalmente, estos residuos contactan con las unidades de desoxirribosa, mediante interacciones hidrófobas, de van der Waals y puentes de hidrógeno (Figura 1.11d) (Otting *et al.*, 1990; Billeter *et al.*, 1993; Hirsch y Aggarwal, 1995; Gruschus *et al.*, 1997, 1999).

Otras interacciones de relevancia involucran la posición 2 del polipéptido que, al igual que en el caso de los residuos 3 y 5, incorpora un residuo de lisina

o arginina en la mayoría de los casos. Algunas estructuras cristalográficas han puesto de manifiesto la participación de esta cadena en contactos con el surco menor de la doble hélice, si bien esto no puede considerarse un rasgo general de los complejos homeodominio/ADN (Hovde *et al.*, 2001; Longo *et al.*, 2007).

Además de las posiciones conservadas 3 y 5, otras altamente variables, tales como 4, 6, 7 y 8, parecen contribuir a la estabilización del complejo proteína/ADN mediante contactos con el esqueleto azúcar-fosfato (Otting *et al.*, 1990; Billeter *et al.*, 1993; Hirsch y Aggarwal, 1995; Gruschus *et al.*, 1997, 1999; Fraenkel y Pabo, 1998).

Por último, es preciso mencionar que, a pesar de su indudable relevancia funcional, ciertos estudios muestran que el brazo N-terminal no es esencial para la formación del complejo (Percival-Smith *et al.*, 1990; Qian *et al.*, 1994b; Fabbro *et al.*, 1996; Dragan *et al.*, 2006).

c) Lazo I-II: Finalmente, el lazo I-II, así como los primeros residuos de la hélice II, interaccionan con el esqueleto azúcar-fosfato del ADN, en el extremo 3' de la secuencia de reconocimiento, a través, principalmente, de los residuos 25 y 31 (Figura 1.11e). Dichas posiciones incorporan tirosina y arginina, respectivamente, en la mayor parte de homeodominios (Kissinger *et al.*, 1990; Otting *et al.*, 1990; Billeter *et al.*, 1993; Hirsch y Aggarwal, 1995; Grant *et al.*, 2000). Los contactos observados en esta región del complejo son importantes para conferir afinidad a la asociación homeodominio-ADN. Adicionalmente, se ha sugerido que otros residuos próximos a estas posiciones podrían ser importantes para un correcto posicionamiento de la hélice III, paso esencial en la formación del complejo (Sato *et al.*, 2004).

Es de destacar que las hélices I y II no participan directamente en el proceso de reconocimiento, salvo a través de la ya mencionada posición 31. Consecuentemente, en algunos casos es posible variar notablemente sus secuencias de aminoácidos sin que esto afecte a la funcionalidad del homeodominio (Shang *et al.*, 1994). No obstante, es preciso mencionar que mutaciones que alteren la red de interacciones establecida entre el lazo I-II, la hélice I y la hélice II, podrían alterar el proceso de ajuste inducido (Stollar *et al.*, 2003), así como modificar levemente la especificidad de secuencia del polipéptido (Shang *et al.*, 1994; Isaac *et al.*, 1995).

1.2.4. Termodinámica de unión de homeodominios a ADN

Durante los últimos veinte años se ha realizado un esfuerzo notable por comprender el mecanismo de acción de los homeodominios. Evidentemente, dicho objetivo requiere no sólo un conocimiento de las características estruc-

turales de la asociación homeodominio-ADN, sino también una comprensión de los detalles termodinámicos del proceso. La indudable relevancia de este aspecto de la interacción queda patente en el elevado número de estudios dedicados a esclarecerlo.

Como suele ser típico en proteínas que se unen de manera específica a ADN (Maerkl y Quake, 2007), los homeodominios se caracterizan por presentar valores altos de constante de afinidad (K_b), que son raramente inferiores a 10^6 y en algunos casos cercanos a 10^{10} M^{-1} (Tabla 1.3). Éstos superan claramente a los observados para otros procesos de asociación de relevancia biológica, como son las uniones proteína-proteína o enzima-sustrato (K_b típicamente entre 10^3 y 10^6 M^{-1} [Fersht, 1999; Mathews *et al.*, 2000; Di

TABLA 1.3: Datos termodinámicos de unión a ADN de algunos homeodominios representativos^a

Homeodominio	K_b (M^{-1})	ΔG_b (kcal/mol) ^b	ΔH_b (kcal/mol)	$-T\Delta S_b$ (kcal/mol)	$\Delta C_{p,b}$ (kcal/[mol·K])	Tampón ^c	Refs. ^h
Antennapedia	$8,3 \cdot 10^7$ a $6,3 \cdot 10^8$ ^d	-10,6 a -11,8 ^d	-	-	-	Q	1, 2, 3, 4
	$2,9 \cdot 10^8$ ^e	-11,6 ^e	-7,2 ^e	-4,4 ^e	-	R	5
	$5,0 \cdot 10^9$ ^f	-13,0 ^f	-2,9 ^f	-10,1 ^f	-0,30 ^f	E	6
Engrailed	$6,7 \cdot 10^8$ ^d	-11,8 ^d	-	-	-	S	7
	$1,3 \cdot 10^{10}$ ^d	-13,8 ^d	-	-	-	T	8
	$1,3 \cdot 10^8$ ^f	-10,8 ^f	-4,5 ^f	-6,3 ^f	-0,23 ^f	E	6
<i>fushi tarazu</i>	$7,7 \cdot 10^8$ a $1,2 \cdot 10^9$ ^d	-11,9 a -12,2 ^d	-	-	-	Q	9, 3
MATα2	$2,5 \cdot 10^7$ ^g	-10,1 ^g	-19,3 ^g	9,2 ^g	-0,53 ^g	U	10
	$7,9 \cdot 10^8$ ^f	-11,9 ^f	-6,2 ^f	-5,7 ^f	-0,30 ^f	E	6
NKX2.5	$4,9 \cdot 10^8$	-11,8	-10,3	-1,5	-0,18	M	11
TTF-1	$4,5 \cdot 10^9$ ^d	-12,9 ^d	-	-	-	Q	12
	$7,5 \cdot 10^7$	-10,7	-9,3	-1,4	-	V	13
vnd (NK-2)	$1,2 \cdot 10^8$	-11,0	-6,6	-4,4	0,25	W	14
	$1,3 \cdot 10^9$ ^f	-12,3 ^f	-5,7 ^f	-6,5 ^f	-0,38 ^f	E	6

^a Salvo que se indique con los superíndices *d*, *e*, *f* o *g*, los valores fueron derivados de experimentos de valoración empleando microcalorimetría (ITC).

^b Derivado de K_b a partir de la expresión $\Delta G_b = -RT \ln K_b$.

^c Los tampones usados y las temperaturas de los ensayos son los siguientes: E, 20 mM acetato pH 5,0, 100 mM NaCl, 20 °C; M, 20 mM MOPS pH 7,4, 100 mM NaCl, 25 °C; Q, 20 mM tris-HCl pH 7,6, 75 mM KCl, 20 °C; R, 10 mM HEPES pH 7,4, 150 mM NaCl, 25 °C; S, 25 mM HEPES pH 7,6, 100 mM KCl, 20 °C; T, 10 mM tris-HCl pH 7,5, 50 mM NaCl, 20 °C; U, 20 mM fosfato pH 7,0, 100 mM NaCl, 25 °C; V, 20 mM acetato pH 5,0, 100 mM KCl, 25 °C; W, 50 mM fosfato pH 7,4, 25 °C.

^d Valores derivados de experimentos de retardo en gel (EMSA).

^e Valores derivados de experimentos de resonancia de plasmones superficiales (SPR).

^f Valores derivados de experimentos de anisotropía de fluorescencia.

^g Valores derivados de experimentos de calorimetría diferencial de barrido (DSC).

^h Referencias: 1, Müller *et al.*, 1988; 2, Affolter *et al.*, 1990; 3, Qian *et al.*, 1994a; 4, Qian *et al.*, 1994b; 5, Seimiya y Kurosawa, 1996; 6, Dragan *et al.*, 2006; 7, Kissinger *et al.*, 1990; 8, Ades y Sauer, 1994; 9, Percival-Smith *et al.*, 1990; 10, Carra y Privalov, 1997; 11, Fodor *et al.*, 2005; 12, Fabbro *et al.*, 1996; 13, Del Vecchio *et al.*, 2008; 14, Gonzalez *et al.*, 2001.

Nardo *et al.*, 2003; Jemth y Gianni, 2007]). Si bien nos movemos siempre en valores de K_b elevados, resulta sorprendente la variabilidad existente al pasar de un homeodominio a otro (hasta cuatro órdenes de magnitud), especialmente si tenemos en cuenta que todos los homeodominios exhiben modos de unión a su diana biológica muy similares.

En términos de energía libre de unión (ΔG_b), los complejos homeodominio/ADN presentan estabilidades comprendidas entre $-8,2$ y $-13,8$ kcal/mol. Esta dispersión de valores resulta de una variabilidad todavía más pronunciada en los componentes entálpico (ΔH_b), entre $-2,9$ y $-19,3$ kcal/mol, y entrópico (ΔS_b), con términos $-T\Delta S_b$ entre $-10,1$ y $9,2$ kcal/mol (Carra y Privalov, 1997; Gonzalez *et al.*, 2001; Fodor *et al.*, 2005; Dragan *et al.*, 2006; Del Vecchio *et al.*, 2008). Dentro de estos extremos, las reacciones de asociación homeodominio-ADN se suelen caracterizar por estar controladas entálpicamente y, al mismo tiempo, por presentar entropías pequeñas y que no se oponen a la unión ($-T\Delta S_b$ en torno a -3 kcal/mol). La entalpía favorable es típica de proteínas que se unen de forma específica al surco mayor (Privalov *et al.*, 1999, 2007). Por su parte, valores bajos y favorables de ΔS_b acostumbran a observarse en proteínas que se unen de manera no específica al surco menor (Privalov *et al.*, 1999, 2007). Conviene destacar que, pese a los bajos valores que suele tomar, ΔS_b resulta de una compensación entre grandes términos entrópicos procedentes, por un lado, de la pérdida de grados de libertad de las especies implicadas y, por otro lado, de la liberación de moléculas del disolvente al disminuir la superficie expuesta en el complejo. Todas estas características termodinámicas son consistentes, a grandes rasgos, con el modelo estructural de asociación homeodominio-ADN descrito en el apartado 1.2.2.

Resulta de especial interés, en complejos proteína/ADN, el análisis de las variaciones en la capacidad calorífica asociadas a la unión ($\Delta C_{p,b}$). De hecho, como propusieron Spolar y Record (Spolar y Record, 1994), valores de $\Delta C_{p,b}$ grandes y negativos son habituales en asociaciones proteína-ADN específicas de secuencia y, además, constituyen la firma termodinámica más característica del ajuste inducido que, con frecuencia, acompaña a la formación de estos tipos de complejos. En homeodominios, dicho parámetro se encuentra habitualmente entre $-0,53$ y $-0,18$ kcal/(mol·K) (Carra y Privalov, 1997; Fodor *et al.*, 2005; Dragan *et al.*, 2006), con la excepción de una única medición llevada a cabo sobre vnd (NK-2), que reveló un $\Delta C_{p,b}$ de $0,25$ kcal/(mol·K) (Gonzalez *et al.*, 2001). Se trata, por lo tanto, de valores quizá pequeños para considerarlos indicativos de este modelo de unión, pese a las numerosas evidencias estructurales que apuntan en ese sentido, y sobre las que profundizaremos más adelante.

Teniendo en cuenta el carácter marcadamente polar de la interacción homeodominio-ADN, no resulta sorprendente que la estabilidad de estos

complejos sea tremendamente sensible a condiciones del medio tales como la fuerza iónica. En general, los datos experimentales disponibles son consistentes con una clara desestabilización de los complejos al aumentar la concentración de sal en el medio (Figura 1.12). A modo de ejemplo, para el homeodominio de MAT α 2 se observa una bajada de 30 veces en la constante de asociación (K_b) al pasar de 50 a 200 mM NaCl (Carra y Privalov, 1997). En el caso de Engrailed, el descenso es todavía más pronunciado: 100 veces al pasar de 140 a 340 mM NaF (Dragan *et al.*, 2006). Finalmente, para los homeodominios de TTF-1 y Antennapedia, se han descrito caídas en K_b de 3000 veces (al pasar de 50 a 400 mM KCl [Del Vecchio *et al.*, 2008]) y 5000 veces (al pasar de 100 a 340 mM NaCl [Dragan *et al.*, 2006]), respectivamente. Estos comportamientos ponen de manifiesto el papel clave que las interacciones electrostáticas juegan en la unión de homeodominios a ADN.

Los elevados valores de las constantes de afinidad para la asociación de homeodominios a ADN son fruto, presumiblemente, de una profunda optimización de las propiedades de unión de estas proteínas y, por lo tanto, el margen de mejora de las mismas mediante mutaciones no debería ser muy grande. De hecho, apenas se dan casos de sustituciones en homeodominios que conlleven una estabilización de su complejo con ADN, pese a existir varios trabajos en esta línea.

La dificultad para mejorar las propiedades de unión de los homeodominios mediante la introducción de mutaciones, así como el descenso de la constante de afinidad al aumentar la fuerza iónica, contrastan con los efectos

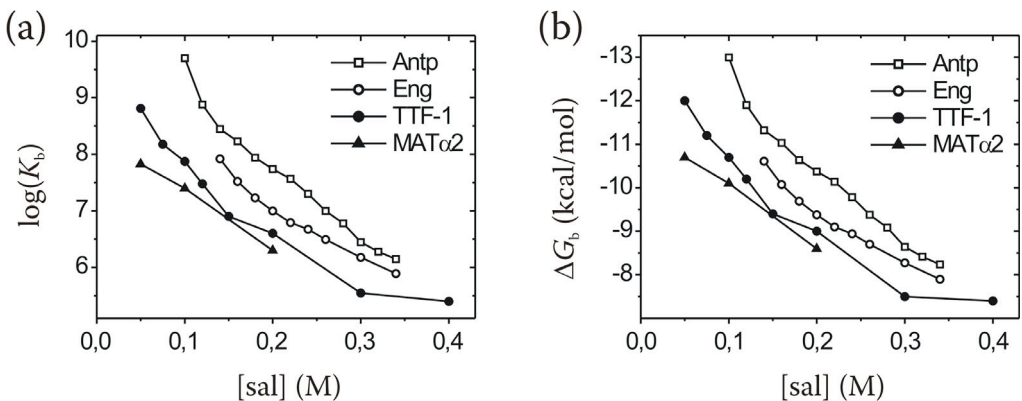


FIGURA 1.12: **Disminución de la afinidad por ADN con la concentración de sal**, observada para los homeodominios de Antennapedia (Dragan *et al.*, 2006), Engrailed (Dragan *et al.*, 2006), TTF-1 (Del Vecchio *et al.*, 2008) y MAT α 2 (Carra y Privalov, 1997). (a) Dependencia de $\log(K_b)$ con [sal]. (b) Dependencia de ΔG_b con [sal]. La sal utilizada fue NaCl para Antennapedia y MAT α 2, NaF para Engrailed, y KCl para TTF-1.

que los mismos procedimientos producen en su estabilidad (apartado 1.1.3) y sugieren que estabilidad y función podrían estar relacionadas de algún modo. De hecho, características de estas proteínas claves en el reconocimiento de ADN, tales como la existencia de regiones desordenadas en los extremos N- y C-terminal, así como una alta carga neta positiva, están en el origen de su baja estabilidad.

IMPLICACIONES TERMODINÁMICAS DEL AJUSTE INDUCIDO

Pese a que los $\Delta C_{p,b}$ típicos de los homeodominios no constituyen pruebas claras de la existencia del ajuste inducido, en apartados anteriores hemos citado las numerosas evidencias estructurales que apuntan a este modelo de reconocimiento de ADN. A continuación vamos a profundizar sobre ellas, así como sobre sus influencias en los parámetros termodinámicos de unión:

a) *Estructuración de la hélice III-b sobre el surco mayor*: La elongación de la hélice III como consecuencia de la unión a ADN es muy evidente en homeodominios como los de vnd (NK-2), TTF-1, MAT α 2, NKX2.5 o *fushi tarazu*, y ha sido puesta de manifiesto por diversos estudios de dicroísmo circular (Carra y Privalov, 1997; Del Vecchio *et al.*, 2008) o RMN (Qian *et al.*, 1994a; Tsao *et al.*, 1994; Fausti *et al.*, 2001; Fodor *et al.*, 2005). No obstante, el efecto es menos pronunciado en casos como los de Engrailed o Antennapedia (Dragan *et al.*, 2006), ya que estos homeodominios parten de una hélice III-b más estructurada en el estado libre (Güntert *et al.*, 1991; Tell *et al.*, 1999; Stollar *et al.*, 2003). Por lo tanto, la magnitud de esta aportación al ajuste inducido puede presentar variaciones significativas entre diferentes homeodominios.

Ciertos autores han propuesto que la estructuración de la hélice III-b durante la asociación a ADN tiene un efecto positivo sobre K_b (Qian *et al.*, 1994a; Tsao *et al.*, 1994; Weiler *et al.*, 1998). La propia generación de hélice α constituye una importante aportación favorable a ΔH_b procedente de las nuevas interacciones intrahélice formadas. No obstante, esta contribución resultaría contrarrestada por la inevitable pérdida de entropía conformacional asociada (Dragan *et al.*, 2006). La explicación al efecto positivo observado en K_b debe buscarse, por tanto, en otro fenómeno. Así pues, la flexibilidad intrínseca de la hélice III-b permitiría una mayor adaptabilidad a la diana de reconocimiento, posibilitando la formación de interacciones entálpicas más eficientes con el ADN y arrojando un balance neto favorable a la asociación. Por otro lado, se ha comprobado que esta propiedad permite a ciertos homeodominios, como es el caso de TTF-1 (Tell *et al.*, 1999), reconocer

distintas secuencias de reconocimiento con afinidades parecidas y, por lo tanto, tendría una incidencia sobre la especificidad de la interacción, incrementando la versatilidad del polipéptido.

b) Inmovilización del brazo N-terminal sobre el surco menor: El evento quizá más llamativo en el proceso de asociación homeodominio-ADN es la transición que experimenta el brazo N-terminal, desde una completa desestructuración en el estado libre, a la adopción de una conformación extendida y perfectamente definida a lo largo del surco menor en el estado asociado.

Varios estudios han cuantificado la aportación del brazo N-terminal a la estabilidad de los complejos homeodominio/ADN. Así pues, se han descrito caídas de 10 veces en K_b en Antennapedia o TTF-1 (Qian *et al.*, 1994b; Fabbro *et al.*, 1996; Dragan *et al.*, 2006), y de hasta 100 veces en *fushi tarazu* o vnd (NK-2) (Percival-Smith *et al.*, 1990; Dragan *et al.*, 2006), asociadas a la delección mutagénica de los seis primeros residuos de cada homeodominio. La disección termodinámica de esta notable contribución ha sido proporcionada, para los homeodominios de Antennapedia y vnd (NK-2), por el grupo de Dragan y Privalov (Dragan *et al.*, 2006). Así pues, la unión del brazo N-terminal está conducida entálpicamente, pero con un término entrópico también favorable. Se ha comprobado que su eliminación no cambia la dependencia de la constante de afinidad respecto a la fuerza iónica. Este resultado es ciertamente sorprendente teniendo en cuenta el carácter policatiónico del brazo N-terminal, y suscita dudas sobre la fuerza dominante en su interacción con el surco menor. Tomadas de un modo global, todas estas observaciones resaltan el peso de los contactos específicos (puentes de hidrógeno y contactos de van der Waals) establecidos por esta región del homeodominio con las bases del ADN. Por último, el aporte del brazo N-terminal a $\Delta C_{p,b}$ es de unas $-0,10$ kcal/(mol·K), lo que, ciñéndonos a las propuestas de Spolar y Record (Spolar y Record, 1994), es indicativo de cierta contribución al ajuste inducido.

c) Compactación global del plegamiento: Pese a que las dos zonas que más claramente se estructuran al unir ADN son el brazo N-terminal y la hélice III-b, el resto de las regiones del homeodominio, incluido el núcleo hidrófobo, podrían, en principio, jugar un papel en el ajuste inducido. La mayor o menor movilidad de dichas regiones podría modular la adaptabilidad de las zonas que contactan la diana de reconocimiento. La pérdida de entropía conformacional asociada se compensaría con un balance entálpico más favorable, debido a una optimización de las interacciones. Un ejemplo de este fenómeno podría venir dado por la red de interacciones, centrada en L26, descrita por el grupo de Fersht para el homeodominio de Engrailed, mediante difracción de rayos X (Stollar *et al.*, 2003). Las distintas conformaciones que pueden adoptar las cadenas laterales implicadas en dicha red se encuentran altamente correlacionadas, e influyen en la posición de residuos clave en la unión a ADN, como son

R31, K46 y Q50. Otro ejemplo mucho más aparente, introducido en el apartado 1.1.3, es el de los homeodominios de Oct-2 POU y ISL-1 (Ippel *et al.*, 1998), cuya estabilidad marginal invita a pensar que se unen al ADN desde un estado parcialmente desplegado. La conformación nativa se alcanzaría sólo como consecuencia del proceso de reconocimiento, lo que constituiría un caso extremo de ajuste inducido.

1.2.5. Unión inespecífica de homeodominios. Relevancia funcional

El reconocimiento de alta afinidad entre los homeodominios y el ADN es siempre específico de secuencia (Gehring *et al.*, 1994a; Wolberger, 1996). No obstante, estos motivos estructurales pueden también llevar a cabo una unión inespecífica (Affolter *et al.*, 1990; Carra y Privalov, 1997; Iwahara y Clore, 2006; Iwahara *et al.*, 2006; Del Vecchio *et al.*, 2008), conducida principalmente por interacciones electrostáticas entre los residuos positivos del homeodominio y los fosfatos del ADN. Desde un punto de vista estructural, diversos estudios muestran que este tipo de unión involucra la misma superficie del polipéptido en su interacción con el ADN que en el caso de la asociación específica (Iwahara y Clore, 2006; Iwahara *et al.*, 2006; Givaty y Levy, 2009; Vuzman *et al.*, 2010). Asimismo, la geometría de ambos tipos de complejo es básicamente idéntica, es decir, con la hélice III del homeodominio a lo largo del surco mayor de la doble hélice y el brazo N-terminal a lo largo del surco menor (Iwahara *et al.*, 2006; Vuzman *et al.*, 2010). Ello implica que el proceso de asociación va acompañado de algunos de los cambios conformacionales característicos de la unión específica, tales como la estructuración parcial del segmento N-terminal (Iwahara y Clore, 2006; Iwahara *et al.*, 2006) o de la hélice III-b (Carra y Privalov, 1997; de la Mata *et al.*, 2002; Fraenkel *et al.*, 1998; Tell *et al.*, 1999), si bien éstos son menos pronunciados. En lo referente a interacciones, se ha observado que, en los modos de asociación inespecíficos, el polipéptido reduce los contactos con las bases y ciertas argininas importantes en la unión gozan de mayor movilidad, lo que podría ser importante para el desplazamiento del homeodominio a lo largo de la doble hebra (Iwahara *et al.*, 2006). Como era de esperar, dichos complejos suelen tener también estabilidades sensiblemente inferiores a las de los específicos (unas 200 veces en el caso del homeodominio de Antennapedia [Affolter *et al.*, 1990] o de HOXD9 [Iwahara y Clore, 2006]). A pesar de ello, no pueden considerarse en modo alguno despreciables, especialmente teniendo en cuenta el tremendo exceso de secuencias de ADN no específicas existente en el interior celular.

Lejos de ser un hecho anecdótico, la unión inespecífica de homeodominios a ADN tiene una clara relevancia funcional, incrementando de forma

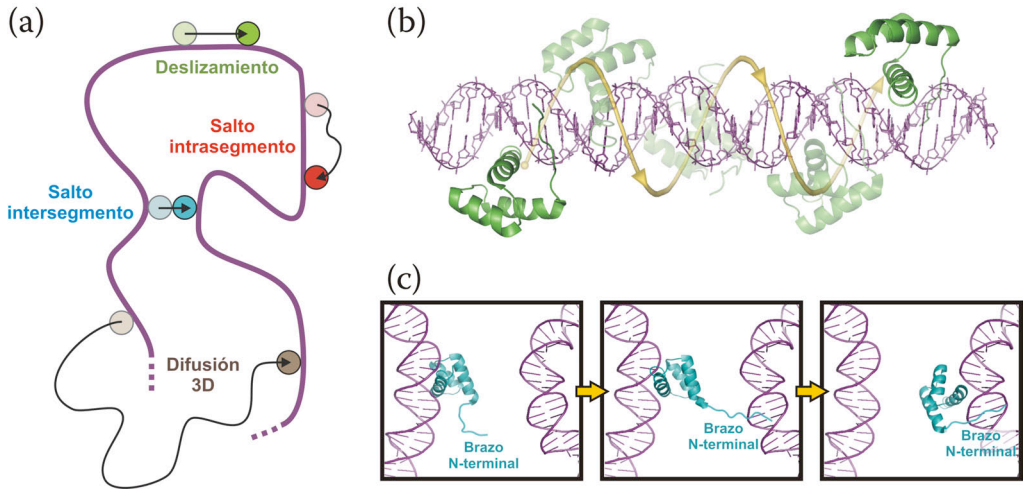


FIGURA 1.13: Los procesos de asociación no específica homeodominio-ADN resultan esenciales en la búsqueda y localización de la secuencia diana. (a) Ilustración esquemática de los procesos de deslizamiento (*sliding*) (verde), saltos intrasegmento (*hopping*) (rojo), saltos intersegmento (azul) y difusión tridimensional (marrón). (b) Detalle del mecanismo de deslizamiento, en el que el homeodominio sigue una trayectoria helicoidal a lo largo de los surcos del ADN. (c) Representación del mecanismo propuesto de saltos intersegmento favorecido por el brazo N-terminal.

significativa la cinética con la cual la proteína localiza su secuencia diana, es decir, acelerando el proceso de búsqueda. Así pues, el primer paso en la acción de los homeodominios implicaría, en la mayoría de los casos, la unión no específica del polipéptido a ADN, seguido de un desplazamiento a través del mismo mediante tres mecanismos generalizados en proteínas de unión a ADN (Berg *et al.*, 1981; Vuzman *et al.*, 2010): deslizamiento sobre el ácido nucleico (*sliding*), pequeños saltos intrasegmento (*hopping*), y saltos intersegmento (Figura 1.13a). Mediante el deslizamiento, se consigue restringir la búsqueda de la secuencia de reconocimiento a una sola dimensión, con la consecuente mejora cinética respecto a lo que sería una búsqueda en tres dimensiones. Los dos tipos de desplazamiento mediante saltos permitirían, por su parte, sortear obstáculos (tales como otras proteínas unidas al ADN) en el proceso de búsqueda. Los tres mecanismos se verían facilitados por la menor estabilidad y mayor carácter dinámico característicos de las uniones no específicas.

Por lo que respecta al mecanismo de deslizamiento, el proceso tendría lugar sin desviaciones evidentes respecto a la geometría típica de los complejos específicos, lo que implica un desplazamiento siguiendo la trayectoria helicoidal marcada por la forma del ADN (Figura 1.13b) (Iwahara *et al.*, 2006;

Givaty y Levy, 2009). En cambio, los mecanismos que involucran saltos parecen requerir la formación de complejos transitorios cuya geometría sí difiere de la observada para complejos específicos (Givaty y Levy, 2009). Así por ejemplo, se ha propuesto que el brazo N-terminal podría facilitar los saltos intersegmento. De acuerdo con esta hipótesis, el homeodominio se uniría al segmento de partida mediante la hélice III, interaccionando al mismo tiempo con el segmento de destino a través del brazo N-terminal, o viceversa (Figura 1.13c) (Vuzman *et al.*, 2010).

En resumen, la función de los homeodominios requiere, en todos los casos, la unión específica del polipéptido a su secuencia de ADN diana. Sin embargo, dicho estado sólo se alcanza tras la formación de una serie de complejos no específicos de carácter transitorio cuya relevancia biológica es incuestionable.

1.3

PATOLOGÍAS HUMANAS RELACIONADAS CON LA DISFUNCIÓN DE HOMEODOMINIOS

Hasta la fecha hay descritas numerosas mutaciones en genes homeobox que dan lugar a enfermedades humanas hereditarias (D'Elia *et al.*, 2001). Un gran número de ellas son puntuales, lo que nos da ocasión de entender la importancia del residuo involucrado sobre la función biológica de la proteína. En otros casos, mutaciones somáticas de estos genes que originan enfermedades adquiridas, como el cáncer (D'Elia *et al.*, 2001), han sido asociadas a alteraciones más severas como deleciones o translocaciones (van Oostveen *et al.*, 1999).

Las enfermedades humanas causadas por mutaciones puntuales en genes homeobox se caracterizan por causar defectos durante el desarrollo que se traducen en fenotipos bien definidos. En varios casos, mutaciones en un único gen homeobox resultan en malformaciones en varios órganos, lo que indica que ese gen se expresa en múltiples tipos celulares (D'Elia *et al.*, 2001).

Dentro de los genes que codifican homeoproteínas, el propio homeobox resulta ser un sitio de alta incidencia de mutaciones asociadas a patologías. De hecho, estas sustituciones resultan ser mucho más frecuentes dentro del homeobox que en el resto del gen. Esto puede deberse a que el homeodominio requiere de una estructura e interacciones muy precisas para realizar su función de unión a ADN, con lo que el cambio de un residuo puede afectarle en mayor medida que a otros dominios de la proteína que gocen de mayor tolerancia estructural y funcional.

La mayor parte de las mutaciones de este tipo que se producen en homeodominios tiene una herencia dominante, lo que en muchos casos se debe a haploinsuficiencia (Hu *et al.*, 1998; Gage *et al.*, 1999). Esto implica que no es suficiente la expresión del alelo no mutado para llevar a cabo con éxito la función del gen. Por otro lado, también se ha descrito algún caso en que la mutación da lugar a un homeodominio que se une con mayor afinidad al ADN diana, lo que también se traduce en enfermedad (Jabs *et al.*, 1993; Ma *et al.*, 1996). Ambas observaciones sugieren que la cantidad de homeoproteínas producidas, así como la estabilidad de los complejos formados con su dianas

biológicas, son parámetros que deben mantenerse dentro de unos límites muy bien definidos. Curiosamente, en ocasiones una mutación pone en marcha mecanismos de dominancia negativa, de forma que el producto del alelo mutante impide la acción del producto del alelo normal (D'Elia *et al.*, 2001; Radovick *et al.*, 1992). Lógicamente, este tipo de sustituciones tiene un efecto especialmente acusado sobre el fenotipo del paciente afectado.

Si bien se han descrito mutaciones puntuales asociadas a patologías a lo largo de toda la secuencia del homeodominio, existen tres localizaciones de alta incidencia: el brazo N-terminal (especialmente la arginina 5), la hélice III (especialmente las posiciones 52 y 53, argininas en casi todos los casos), y el residuo 31 (normalmente arginina). Como hemos visto en el apartado 1.2.3, estas cuatro argininas interaccionan con los fosfatos del esqueleto de ADN, y parecen ser relevantes en la estabilización del complejo proteína/ADN. El hecho de que sean estas posiciones los sitios de mayor incidencia de mutaciones asociadas a enfermedades sugiere una asombrosa conclusión: que los contactos arginina-fosfato son más importantes de cara a la funcionalidad del homeodominio que los que se producen con las bases del ADN.

De todas las posiciones afectadas por mutaciones puntuales que dan lugar a las diversas enfermedades, destacamos las siguientes, de mayor o menor importancia, pero relacionadas directamente con la presente tesis:

- **Arginina 31 y ácido glutámico 42:** además de su relevancia para el contacto con el ADN, hemos visto que la arginina 31 tiene un señalado papel estructural al formar un puente salino con el glutámico en posición 42. La importancia de ambos residuos se ve reflejada en el gran número de mutaciones descritas para la posición 31, y en menor número para la 42. A modo de ejemplos, la mutación R31P en el homeodominio de MSX-1 impide el desarrollo de algunos dientes (Vastardis *et al.*, 1996), y la mutación R31H en el homeodominio de PITX2 (RIEG1) afecta a la parte anterior del ojo (síndrome de Rieger) (Kulak *et al.*, 1998). En cuanto al glutámico 42, la mutación E42A del homeodominio de CRX causa distrofia de los conos y bastones de la retina y en consecuencia una pérdida progresiva de la visión (Freund *et al.*, 1997).

- **Prolina 26:** al no ser una posición que contacte con el ADN, las enfermedades asociadas a su alteración resultan en cierto modo sorprendentes, y ponen de manifiesto la importancia de 26 en otras importantes propiedades del homeodominio. Si bien en esta posición es frecuente encontrar un residuo alifático, abundan las secuencias, como POU1F1 y ARX, que incluyen una prolina. En el caso de POU1F1 la sustitución P26S da lugar a hipoplasia de la pituitaria anterior, con la consecuente falta de producción de hormona de crecimiento, prolactina y tirotropina (Pernasetti *et al.*, 1998). Dicha sustitución

apenas afecta a la unión a ADN de POU1F1, con lo que su efecto patogénico ha sido asociado a la incapacidad del homeodominio mutante para establecer contactos interproteína necesarios para la activación de la transcripción. Su modo de herencia es recesivo, en contraste con lo que sucede en el caso del alelo mutante más abundante para esta enfermedad, R58W, que da lugar a dominancia negativa. Otro importante ejemplo es la sustitución P26L en el homeodominio de *ARX*, que causa epilepsia mioclónica acompañada de retraso mental y espasticidad (Scheffer *et al.*, 2002; Strømme *et al.*, 2002). Esta mutación llama especialmente la atención porque, a simple vista, parece conservativa, ya que muchos homeodominios humanos tienen leucina en posición 26. Su efecto patogénico ha sido asociado a la alteración de una importante red de interacciones centrada en el residuo 26 (Stollar *et al.*, 2003), que hemos descrito en el apartado 1.2.4.

- **Residuo en posición 37:** la mutación R37L en el homeodominio de *SHOX* causa discondrosteosis, una displasia esquelética caracterizada por una estatura desproporcionadamente baja (Grigeliene *et al.*, 2000). Si bien el residuo en 37 no contacta con el ADN, en este homeodominio la arginina en esa posición podría formar un puente salino con el glutámico 15 y con el glutámico 33. El efecto de la mutación R37L sobre el fenotipo sugiere que esos puentes salinos son importantes para la estabilidad o bien para el correcto plegamiento del homeodominio de *SHOX*.

- **Arginina 52:** Además de su papel en el reconocimiento de ADN, en el apartado 1.1.2 hemos comentado que la arginina 52 participa en la formación de un puente salino con el glutámico en posición 17. Este puente salino es el más conservado en homeodominios, (está presente en el 75% de las secuencias descritas), lo que sugiere su gran importancia funcional. Por este motivo resulta extraño que no haya descrita ninguna enfermedad causada por una mutación de cambio de sentido en el glutámico 17, mientras que la posición 52 es, como se ha dicho antes, un punto de alta incidencia de mutaciones. Entre las mutaciones de la arginina 52 causantes de enfermedades, cabe destacar la sustitución R52P en *SIX3*, responsable de un tipo de holoprosencefalia (separación del sistema nervioso anterior en dos mitades) (Wallis *et al.*, 1999). Por su parte, la sustitución R52C en *PROP1* causa una deficiencia en la producción de hormonas de la pituitaria (Wu *et al.*, 1998). Asimismo, la misma mutación (R52C) en *PITX2* (*RIEG1*) es otra de las causas descritas del síndrome de Rieger (Perveen *et al.*, 2000). Finalmente, la sustitución R52W en el homeodominio de *CRX* tiene un efecto sobre la retina similar al de la mutación E42A anteriormente mencionada (Swaroop *et al.*, 1999). En algunos de los trabajos citados se muestran ensayos de retardo en gel que revelan menor afinidad por ADN causada por las mutaciones en R52, lo cual se

propone como la causa de las enfermedades descritas (Wu *et al.*, 1998; Swaroop *et al.*, 1999). Sin embargo, la rotura del puente salino 17-52 y su impacto en la estabilidad del homeodominio podría ser también importante en el origen de los fenotipos observados.

- **Arginina 53:** Al igual que R52, R53 es importante en el balance electrostático de la unión del homeodominio al ADN. Sin embargo, a diferencia del primer residuo, R53 establece contactos con el esqueleto del ADN claves para la formación del complejo. La importancia funcional de la posición 53 está subrayada, como ya hemos avanzado, por el hecho de que constituye un punto de alta incidencia de mutaciones asociadas a patologías. A modo de ejemplo, la mutación R53P en PITX2 también da lugar al síndrome de Rieger (Semina *et al.*, 1996), y la mutación R53Q en HLXB9 provoca el síndrome de Currarino (caracterizado por diversas anomalías anorrectales y del sacro) (Hagan *et al.*, 2000). Por último, la mutación R53C en el homeodominio de HESX-1 causa displasia septo-óptica (Dattani *et al.*, 1998), como se explica en la parte de antecedentes y objetivos a continuación.

2 *ANTECEDENTES Y OBJETIVOS*

2.1

ANTECEDENTES DEL GRUPO DE INVESTIGACIÓN: EL HOMEODOMINIO DEL FACTOR DE TRANSCRIPCIÓN HUMANO HESX-1

Hesx-1, identificado en primer lugar como un gen homeobox expresado en células madre embrionarias (ES, *embryonic stem*) de ratón (Thomas y Rathjen, 1992), pertenece a la clase *Anf* (Kazanskaya *et al.*, 1997). Durante etapas tempranas del desarrollo, se expresa inicialmente en varias zonas que posteriormente generarán distintos centros del sistema nervioso anterior (Dattani *et al.*, 1998), y más tarde en el hígado embrionario (Thomas *et al.*, 1995). La falta de actividad de HESX-1 en humanos es el origen de una enfermedad genética conocida como displasia septo-óptica (Dattani *et al.*, 1998, Thomas *et al.*, 2001), caracterizada por defectos en la morfogénesis de la pituitaria, el nervio óptico y ciertas estructuras cerebrales. Esta patología puede estar originada por diversas mutaciones. A modo de ejemplo, hace unos años se identificó la sustitución R53C como la causa de esta pérdida de actividad en algunos pacientes, que mostraban una herencia recesiva de la enfermedad (Dattani *et al.*, 1998). Asimismo, se mostró cómo dicha mutación destruye la capacidad de HESX-1 de unir su ADN diana, si bien no afecta en gran medida a su estructura terciaria.

Nuestro grupo de investigación ha estado en los últimos años involucrado en el estudio de la estructura, estabilidad y reconocimiento de ADN del homeodominio de HESX-1 humano. Como primer paso se resolvió su estructura tridimensional en disolución empleando RMN, lo que supuso la primera descripción estructural de un homeodominio de la clase *Anf* (de la Mata *et al.*, 2002) (Figura 2.1). También se determinó la estructura del mutante R53C, que resultó poseer un plegamiento típico de homeodominio, con sólo algunas diferencias locales respecto al polipéptido salvaje. No obstante, se comprobó una clara disminución de la estabilidad térmica del mutante R53C, de 10 °C. Además, se confirmó que éste presentaba una baja afinidad por ADN, más de tres órdenes de magnitud inferior a la observada para la proteína salvaje. Dicha afinidad podría verse adicionalmente comprometida en el mutante por la formación de un puente disulfuro entre las cisteínas 53 y 24, lo que explicaría su inactividad *in vivo*.

Con estos antecedentes, la presente tesis doctoral pretende profundizar en el conocimiento de los diversos factores que condicionan la estructura, estabilidad y función de los homeodominios, empleando HESX-1 como sistema modelo.

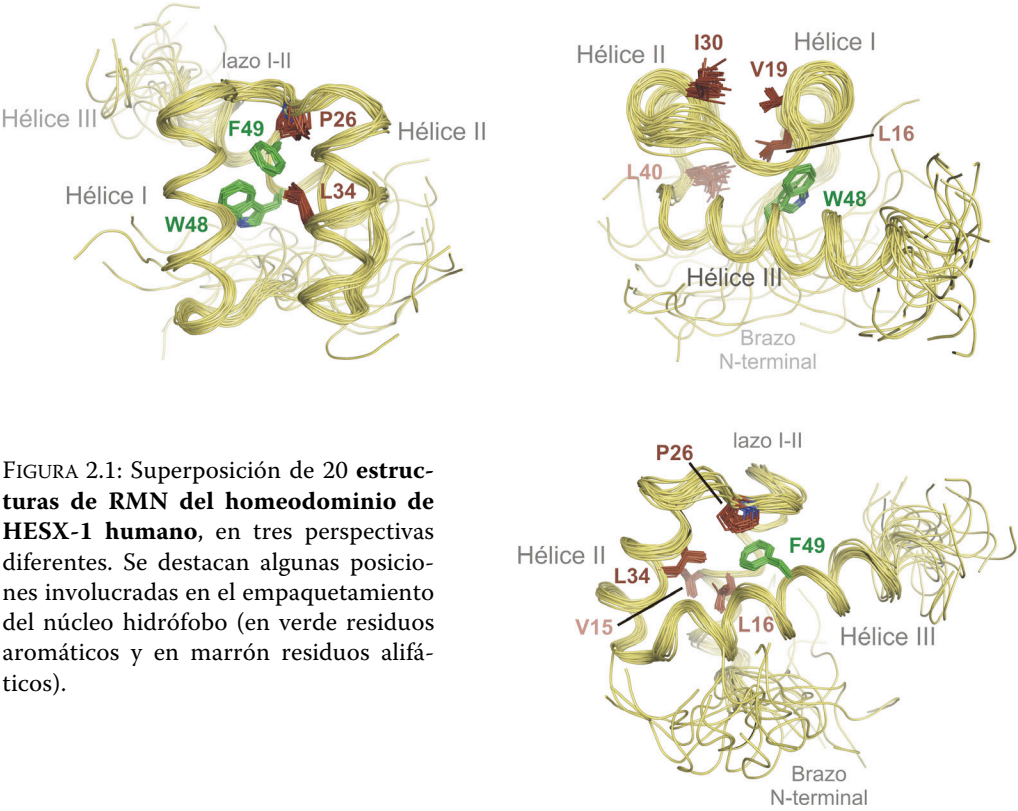


FIGURA 2.1: Superposición de 20 **estructuras de RMN del homeodominio de HESX-1 humano**, en tres perspectivas diferentes. Se destacan algunas posiciones involucradas en el empaquetamiento del núcleo hidrófobo (en verde residuos aromáticos y en marrón residuos alifáticos).

2.2

PLANTEAMIENTO DEL OBJETIVO PRINCIPAL: LA RELACIÓN ENTRE LA ESTABILIDAD Y LA FUNCIÓN DE HOMEODOMINIOS

A lo largo de la introducción hemos hecho un repaso a la información existente sobre secuencia, estructura, estabilidad y función de homeodominios (entendida ésta última como la unión específica a ADN). Las relaciones que se dan entre estas propiedades han sido objeto de numerosos estudios en otros sistemas, lo que ha conducido a varias conclusiones. Así pues, está bien establecido que la estructura de una proteína está dictada por su secuencia de aminoácidos. A su vez, son las características estructurales las que, en combinación con la secuencia, determinan la estabilidad conformacional de una proteína.

Curiosamente, a pesar de su posible relevancia y de ser un tema de debate hoy en día, la correlación existente entre estabilidad y función ha sido objeto de un número de estudios mucho más limitado. En principio, parece claro que la estabilidad determina la fracción de proteína plegada bajo unas condiciones dadas y, por lo tanto, modula su actividad (especialmente en condiciones en las que la estabilidad es marginal, es decir, a temperaturas próximas a la T_m). Sin embargo, más allá de esta relación evidente, estabilidad y función están íntimamente conectadas a través del concepto de flexibilidad conformacional. Esto es así porque, por un lado, una mayor estabilidad conlleva normalmente una menor flexibilidad (Fields, 2001), y por otro lado, la función de una proteína puede mostrarse estrechamente vinculada con su flexibilidad (Ó Fágáin, 1995; Freire, 2001; Teilum *et al.*, 2011).

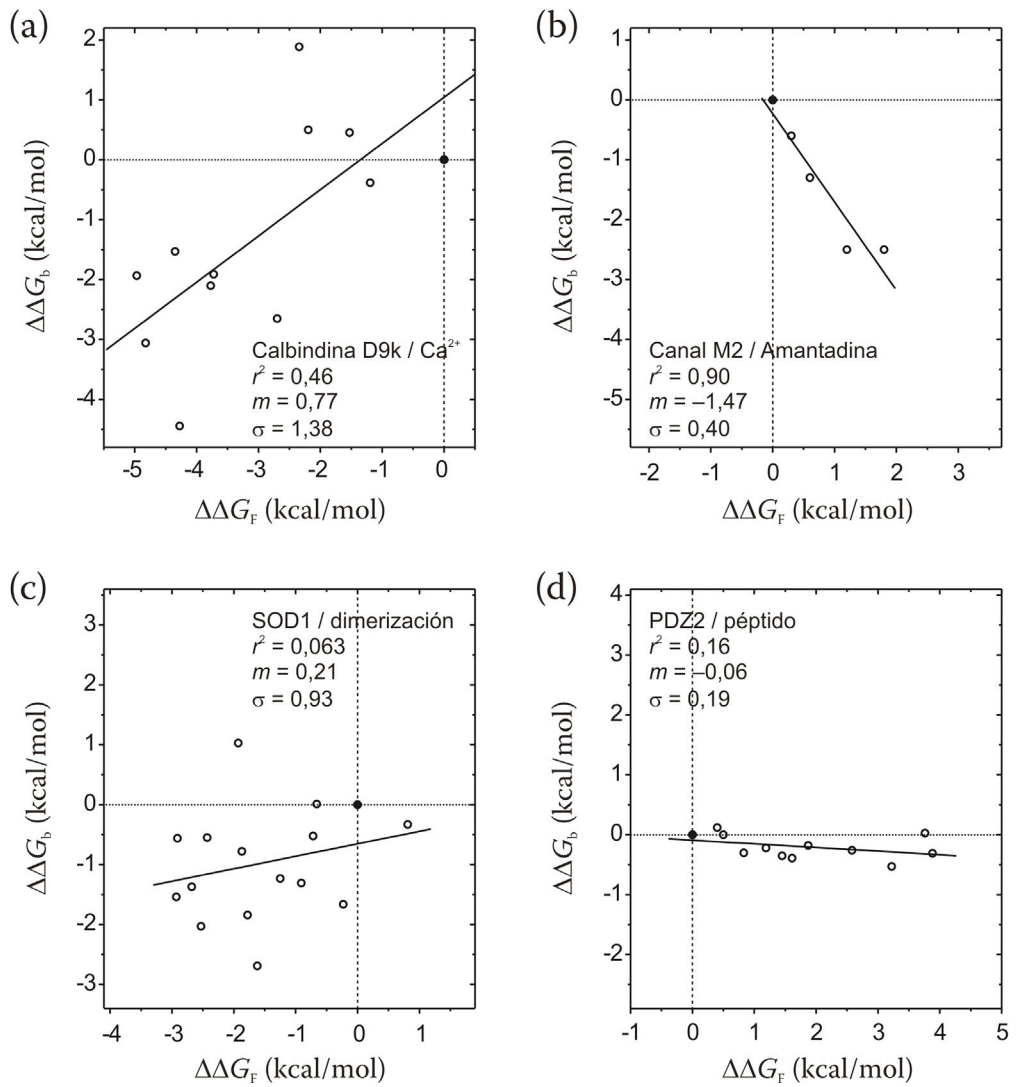
En principio, una aproximación sencilla para analizar experimentalmente la relación entre función y estabilidad sería generar diversas variantes de la proteína objeto de estudio, y someterlas a un análisis detallado de ambas propiedades. Sin embargo, con objeto de acotar el problema y obtener información de interés, dicho estudio debe, en nuestra opinión, considerar tres requisitos previos:

1. En primer lugar, las variantes contempladas no deben presentar modificaciones en zonas de la proteína directamente involucradas en la función.

2. En segundo lugar, en las condiciones de los ensayos, todas las variantes consideradas deben ser estables, presentando fracciones despreciables de proteína desnaturalizada.

3. Finalmente, todas las variantes consideradas deben estar correctamente plegadas.

Tal como se ha mencionado anteriormente, si se contemplan estas condiciones, entraría en juego el grado de flexibilidad de la proteína como concepto básico para entender la conexión entre estabilidad y función. Desafortunadamente, una inspección detallada de la bibliografía revela que los estudios publicados



hasta la fecha consistentes con los requerimientos previamente mencionados son poco numerosos, y pueden englobarse en tres grupos dependiendo de la correlación entre estabilidad y función que revelan:

- **Correlación positiva:** Es lo más habitual en casos de unión proteína-ligando y su explicación es sencilla. Una mayor flexibilidad da lugar, en general, a una peor definición de la superficie de reconocimiento (Fields, 2001; Di Nardo *et al.*, 2003), lo que puede traducirse en interacciones proteína-ligando menos optimizadas. Por otro lado, una mayor flexibilidad implica una mayor penalización entrópica asociada a la compactación que sufre la proteína cuando se forma el complejo (Jemth y Gianni, 2007). De acuerdo con esto, un aumento en la estabilidad de la proteína supondría una reducción de su flexibilidad, y conllevaría así un incremento en su afinidad por el ligando (Figura 2.2a), tal y como se observa en varios trabajos (Julenius *et al.*, 1998; Kragelund *et al.*, 1998; Fields, 2001; Di Nardo *et al.*, 2003; Jemth y Gianni, 2007; Teilum *et al.*, 2011).
- **Correlación negativa:** Surge cuando la flexibilidad de una proteína es requerida para su funcionalidad, como sucede con las enzimas (Fields, 2001), los canales transmembrana (Stouffer *et al.*, 2005), o los casos de alosterismo (Freire, 2001; Travaglini-Allocatelli *et al.*, 2009; Teilum *et al.*, 2011). En estos sistemas, un aumento en la estabilidad haría más rígida la proteína, lo que conllevaría bajadas de actividad (Figura 2.2b).
- **Ausencia de correlación:** Puede ser de dos tipos, según la dispersión de los datos estabilidad-función obtenidos:
 - *Datos dispersos, pobremente ajustables a una recta:* Se observa cuando los cambios en estabilidad causados por las mutaciones vienen acompañados

FIGURA 2.2 (PÁGINA ANTERIOR): **Ejemplos de relaciones estabilidad-función** deducidas a partir del estudio de diversos mutantes de receptores proteicos. (a) La unión de calcio a la calbindina D_{9k} (Julenius *et al.*, 1998; Kragelund *et al.*, 1998) manifiesta una correlación estabilidad-función positiva. (b) Correlación negativa observada para la unión de amantidina al canal de protones M2 del virus de la gripe (Stouffer *et al.*, 2005). (c) La formación del homodímero SOD1 (Lindberg *et al.*, 2005) revela una baja correlación estabilidad-función con alta dispersión de datos. (d) Baja correlación con poca dispersión de datos observada para la unión del dominio PDZ2 de la proteína tirosina fosfatasa-BL a un péptido diana (Jemth y Gianni, 2007). En todas las gráficas y en ambos ejes se ha usado la misma escala para facilitar las comparaciones. El punto correspondiente a la proteína salvaje en cada caso se representa relleno. Los valores $\Delta\Delta G_F = \Delta G_F^{\text{salvaje}} - \Delta G_F^{\text{mutante}}$, con lo que los valores negativos indican menor estabilidad que la proteína salvaje, y los positivos mayor estabilidad. Asimismo, $\Delta\Delta G_b = \Delta G_b^{\text{salvaje}} - \Delta G_b^{\text{mutante}}$, con lo que los valores negativos indican menor afinidad de la unión que para la proteína salvaje, y los positivos mayor afinidad. Las líneas negras representan los ajustes de los datos por mínimos cuadrados. r^2 , coeficiente de correlación elevado al cuadrado. m , pendiente de la recta. σ , desviación típica.

unas veces de aumentos y otras veces de disminuciones de funcionalidad, dependiendo del lugar y la naturaleza de las mutaciones (Figura 2.2c). Se trata de una situación muy poco frecuente en los trabajos publicados (Lindberg *et al.*, 2005; Teilum *et al.*, 2011) y en ocasiones es la consecuencia de un diseño inapropiado de los mutantes.

- *Datos poco dispersos, ajustables a una recta, pero de pendiente cercana a cero*: Se observa cuando mutaciones que afectan a la estabilidad no tienen un efecto apreciable sobre la función (Figura 2.2d). Ejemplos que muestren este comportamiento son incluso menos frecuentes que los descritos en el caso anterior (Jemth y Gianni, 2007).

Si bien se desconoce en cuál de estos grupos encajan los homeodominios, la gran cantidad de información acumulada sobre los mismos, y que hemos resumido en la introducción, nos permite plantear una serie de hipótesis que, en principio, podrían ser comprobadas experimentalmente.

En primer lugar, los homeodominios suelen caracterizarse por presentar estabilidades marginales, con T_{ms} próximas a las temperaturas a las que deben actuar *in vivo*. Este rasgo, compartido por otros tipos de factores de transcripción (Padmanabhan *et al.*, 1997, 1999; Turner *et al.*, 2004; Mason *et al.*, 2006; Mayer *et al.*, 2007; Meier-Andrejszki *et al.*, 2007), podría ser necesario para el correcto desempeño de su función. La estabilidad marginal de los homeodominios tendría como consecuencia la existencia de una proporción apreciable de proteína desnaturalizada. En este escenario, pequeños incrementos o decrementos de estabilidad (por ejemplo, causados por mutaciones) disminuirían o aumentarían aún más dicha proporción, modulando de esta forma la cantidad de proteína nativa, funcionalmente activa, presente en disolución. Tal y como se ha mencionado anteriormente, esta conexión es evidente y, por lo tanto, carece de especial interés.

En segundo lugar, podrían plantearse situaciones en los que la flexibilidad conformacional del homeodominio tuviera implicaciones sobre su función. En el escenario más simple, dicha propiedad elevaría el coste de la inmovilización del homeodominio asociada a la formación del complejo, o podría afectar a la definición de su superficie de interacción con ADN (Figura 2.3a). En estos casos, tal y como se ha explicado anteriormente, sería de esperar la existencia de una correlación positiva entre estabilidad y función. Alternativamente, se podrían concebir situaciones más complejas. En la introducción hemos visto que la asociación homeodominio-ADN constituye un ejemplo de ajuste inducido. Dicho mecanismo podría no implicar de forma exclusiva las partes desestructuradas del polipéptido, sino también el resto de regiones del mismo, incluido el núcleo hidrófobo (Figura 2.3b). En este contexto, la flexibilidad que caracteriza a dichas regiones podría ser esencial para el correcto estableci-

miento de interacciones con el ADN durante el proceso de asociación, determinando una correlación estabilidad-función negativa.

Como puede verse, la relación existente entre estos dos parámetros en homeodominios puede, en principio, revelar detalles sutiles del proceso de asociación, tales como la extensión del mecanismo de ajuste inducido o el papel de las regiones estructuradas en el mismo, difícilmente accesibles a partir de otros datos. En definitiva, la interrelación entre estabilidad y función en estos sistemas representa un asunto de gran complejidad y relevancia, y su análisis constituye el objetivo central de esta tesis. A continuación, se presenta una breve descripción de este objetivo principal y de otros objetivos secundarios, así como de los procedimientos experimentales diseñados a tal efecto.

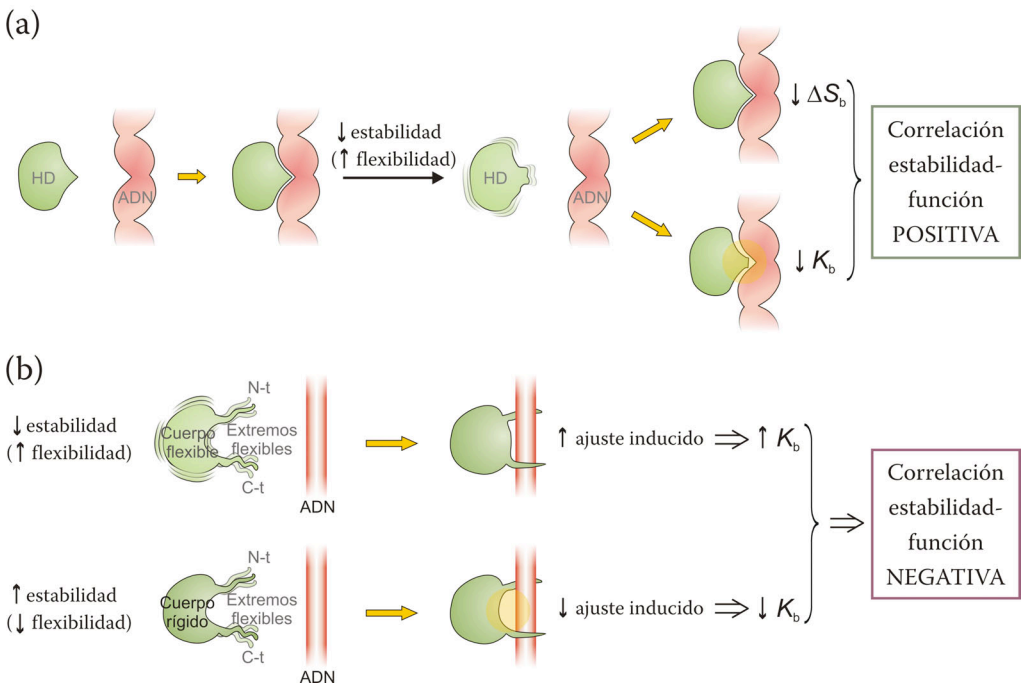


FIGURA 2.3: Se representan esquemáticamente distintos fenómenos que podrían generar una **correlación estabilidad-función positiva o negativa en el homeodominio**. (a) Una reducción en la estabilidad conformacional de la proteína normalmente conlleva un incremento en su flexibilidad. Esto podría determinar un mayor coste entrópico asociado a la formación del complejo y una peor definición de la superficie de unión. Ambas situaciones favorecerían una correlación estabilidad-función positiva. (b) El efecto positivo del ajuste inducido sobre la afinidad de la asociación podría no proceder únicamente de las regiones desordenadas N- y C-terminal del homeodominio, sino implicar también, en cierta medida, la región estructurada. En este caso, la flexibilidad de la misma, cuya magnitud guarda relación con la estabilidad conformacional, podría determinar una correlación estabilidad-función negativa.

2.3

DESARROLLO DEL OBJETIVO PRINCIPAL Y DE LOS OBJETIVOS SECUNDARIOS

2.3.1. Objetivo principal

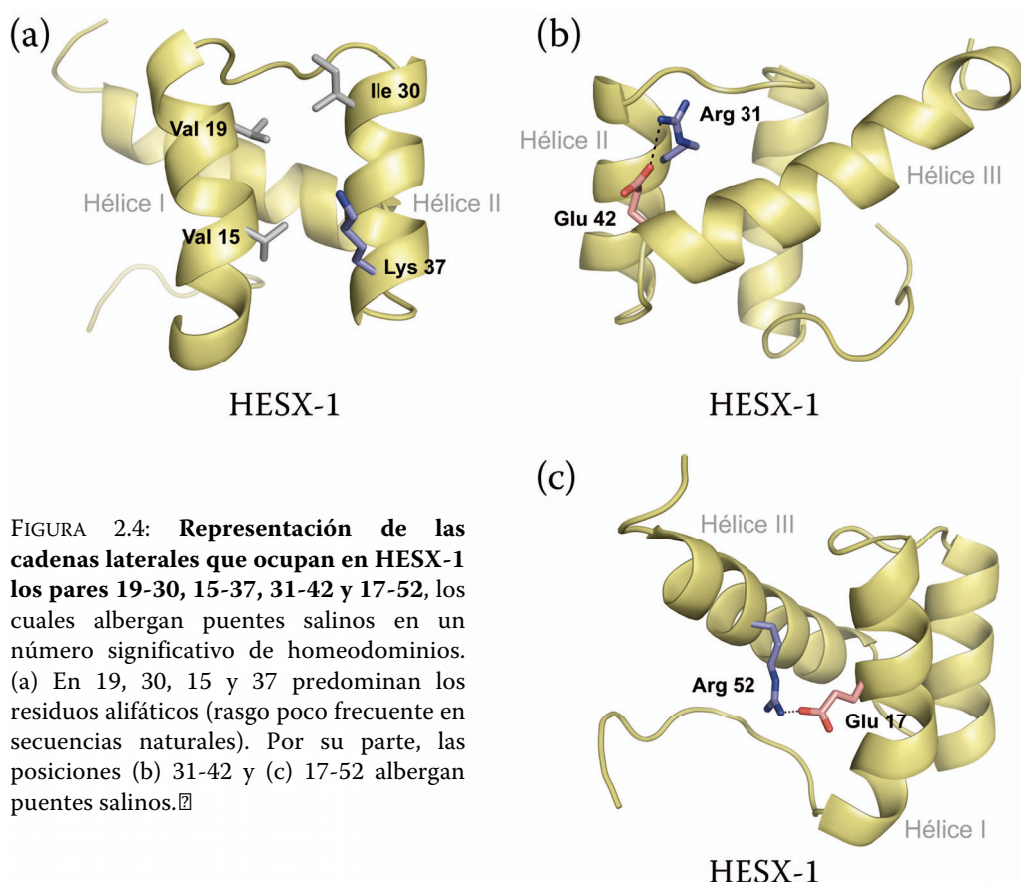
El trabajo recogido en la presente memoria tiene como objetivo principal esclarecer las relaciones existentes entre estabilidad y función de homeodominios. Para ello, usaremos como modelo el dominio de unión a ADN del factor de transcripción humano HESX-1.

El primer paso consiste en generar, mediante mutagénesis dirigida, un conjunto de variantes del homeodominio que abarquen un rango amplio de estabilidades. De acuerdo con los requisitos que hemos apuntado anteriormente, las mutaciones seleccionadas deben además satisfacer las siguientes condiciones:

1. En primer lugar, deben tener un efecto claro sobre la estabilidad del homeodominio, sin afectar posiciones involucradas en la interacción con ADN o próximas a ellas.
2. En segundo lugar, deben generar variantes suficientemente estables, esto es, con una proporción mínima de proteína desplegada a las temperaturas a las que llevemos a cabo los estudios de asociación a ADN.
3. Finalmente, deben ser conservativas. Esta condición asegura, en principio, que todas las variantes compartan un plegamiento común. De este modo, las modificaciones en los parámetros medidos causadas por las mutaciones no podrán relacionarse con cambios acusados en la estructura tridimensional. Para ello, los mutantes serán diseñados de acuerdo con un análisis previo de un extenso número de secuencias de homeodominios.

Con objeto de satisfacer estas tres condiciones, hemos centrado nuestro estudio en las posiciones fronterizas 15, 19 y 30, así como en 37 (algo más expuesta al disolvente). Todas ellas pertenecen a las hélices I y II y, por lo tanto, se localizan en la cara opuesta a la superficie de interacción con ADN. Tal y

como hemos visto en la introducción, en secuencias naturales estas posiciones acostumbran a estar ocupadas por residuos cargados capaces de formar puentes salinos (ver Figuras 1.4a y 1.5a). No obstante, también es frecuente encontrar residuos alifáticos parcialmente expuestos que empaquetan con el núcleo hidrófobo de la proteína. Éste es el caso de HESX-1, que incluye la tétrada V15/V19/I30/K37 (ver Figura 2.4a). Teniendo en cuenta toda esta información, se generaron diversos mutantes del homeodominio natural que incorporan, en posiciones 19, 30, 15 y 37, distintas combinaciones de cadenas laterales polares cargadas, presentes en un buen número de secuencias conocidas. Se trata, por lo tanto, de mutaciones presumiblemente conservativas, que no deberían conllevar cambios estructurales importantes, lo que haría factible, en principio, barrer un amplio abanico de estabilidades, pero manteniendo unos niveles mínimos requeridos para el estudio experimental. Una vez obtenidos los mutantes, se abordó un análisis detallado de su estabilidad. A su



vez, se estudiaron las propiedades de unión a ADN de un conjunto representativo de los mismos.

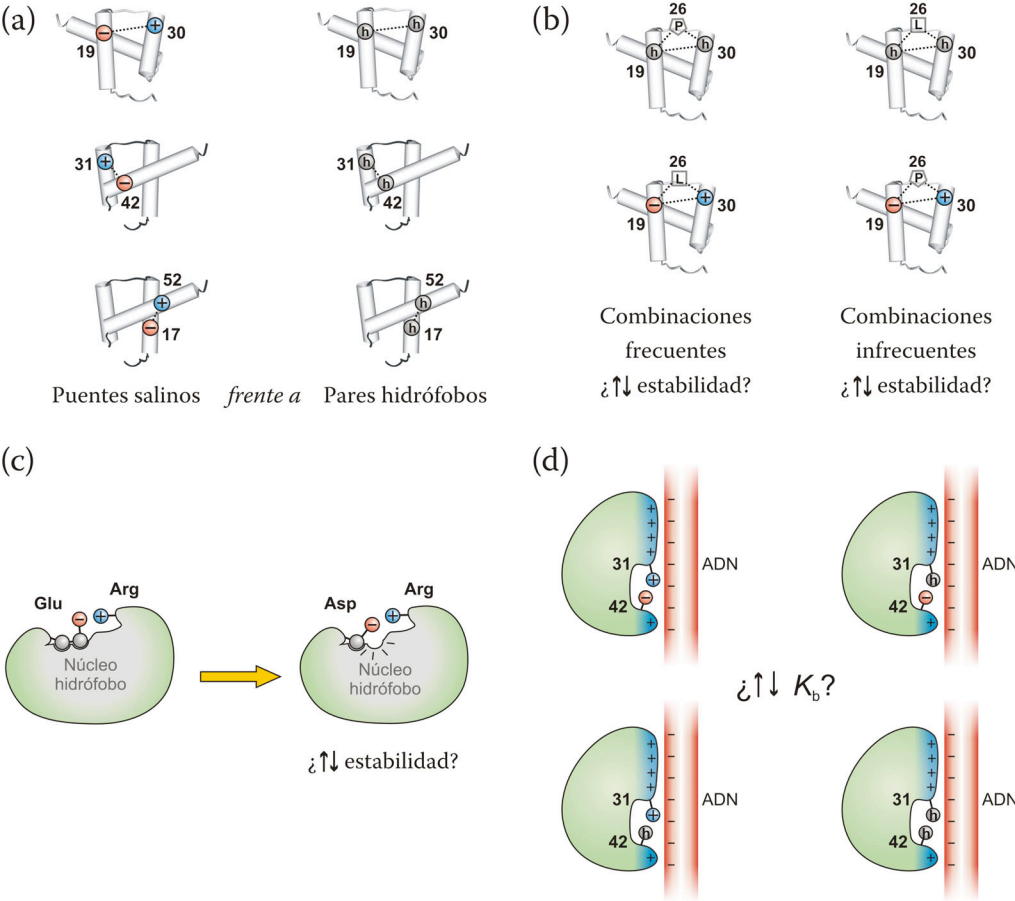


FIGURA 2.5: Ilustración esquemática de los **cuatro objetivos secundarios de la presente tesis**. (a) Comparación del efecto estabilizante aportado por puentes salinos en relación con pares hidrófobos. (b) Las consecuencias de la correlación entre las posiciones 19, 30 y 26 en la estabilidad del homeodominio se examinaron mediante el análisis de distintas combinaciones de residuos en dichas posiciones. *P* significa prolina y *L* leucina. (c) Las cadenas hidrocarbonadas de los residuos involucrados en puentes salinos pueden contribuir a la estabilidad del homeodominio mediante su empaquetamiento con el núcleo hidrófobo. La eliminación de un único grupo metileno mediante mutaciones simples Glu→Asp permite evaluar este efecto. Los círculos grises representan grupos metileno, y los símbolos + y - cargas positivas y negativas, respectivamente. (d) La sustitución de los residuos cargados R31 y E42 por cadenas laterales alifáticas permite evaluar el papel de los primeros en la unión a ADN. Para (a), (b) y (d), los símbolos *h*, + y - se refieren a residuos hidrófobos (V, L o I, o alternativamente Y para las posiciones 17 y 52), arginina y ácido glutámico, respectivamente.

2.3.2. Objetivos secundarios

Las posibilidades abiertas por las variantes generadas inspiraron la consideración de varios objetivos secundarios:

a) En primer lugar, nos planteamos obtener información acerca del efecto estabilizante que tienen los puentes salinos conservados entre las hélices I y II en relación con los pares hidrófobos (Figura 2.5a). Para ello, analizamos las variantes destinadas al objetivo principal, en las que los residuos 19 y 30, valina e isoleucina respectivamente en HESX-1, aparecen sustituidos por aminoácidos cargados, formando puentes salinos de diferentes polaridades. Las variantes con sustituciones adicionales en las posiciones 15 y 37, nos permitieron comprobar el efecto de la formación de redes de puentes salinos que involucran las cuatro posiciones citadas.

Decidimos a su vez extender este estudio a los pares 31-42 y 17-52, ocupados por puentes salinos altamente conservados en la gran mayoría de los homeodominios. HESX-1 incluye en dichas posiciones los residuos R31-E42 y E17-R52 (Figura 2.4, b y c). Por tanto, procedimos a su sustitución por distintas combinaciones hidrófobas observadas en otras secuencias naturales minoritarias.

b) En segundo lugar, decidimos clarificar el origen de la correlación existente entre la naturaleza del par 19-30 y el residuo 26 del lazo I-II (prolina o alifático en la inmensa mayoría de los homeodominios) (Clarke, 1995; Stollar *et al.*, 2003). Con este objetivo, se consideraron variantes adicionales, con puentes salinos o hidrófobos conectando 19-30 en el contexto de prolina o leucina en la posición 26 (Figura 2.5b).

c) En tercer lugar, tal como hemos mencionado anteriormente, las posiciones fronterizas parecen contribuir a la estabilidad conformacional del homeodominio no sólo mediante interacciones electrostáticas, sino a través del empaquetamiento de sus cadenas hidrocarbonadas con el núcleo hidrófobo. Con objeto de cuantificar esta última contribución, obtuvimos dos nuevas variantes en las que se reemplazó un ácido glutámico involucrado en un puente salino por un ácido aspártico (un metileno más corto) (Figura 2.5c). Más concretamente, se compararon las estabilidades conformacionales de variantes que incorporan los pares E19-R30 y D19-R30 por un lado, y R31-E42 y R31-D42 por otro.

d) Finalmente, se analizaron las propiedades de unión a ADN de variantes con mutaciones en el par 31-42, con objeto de esclarecer el papel del mismo en el reconocimiento (Figura 2.5d).

2.3.3. Procedimientos experimentales destinados al abordaje de los objetivos

El trabajo que se detalla en esta tesis conllevó la producción de 33 variantes del homeodominio de HESX-1, incluyendo mutantes simples, dobles, triples y cuádruples. La consecución del objetivo principal y de los cuatro objetivos secundarios implicó un profundo análisis de las variantes basado en las siguientes metodologías experimentales:

- Todas las variantes fueron sujetas a cuidadosos ensayos de desnaturalización térmica y química monitorizados por dicroísmo circular. Los primeros fueron realizados a cinco fuerzas iónicas diferentes para determinar la dependencia de la energía libre de plegamiento (ΔG_F) con la concentración de sal, siguiendo un protocolo similar al descrito por Schmid y colaboradores (Perl *et al.*, 2000; Martin *et al.*, 2002; Wunderlich *et al.*, 2005; Max *et al.*, 2007).
- Mediante métodos de resonancia magnética nuclear (RMN), pudimos obtener una información estructural más precisa, o incluso una estructura detallada en disolución, de aquellos mutantes que, por su peculiar comportamiento, merecieron una consideración especial.
- Los valores de pK_a para los grupos carboxilo de ciertos mutantes seleccionados fueron medidos por espectroscopía de RMN, para profundizar en el papel que los puentes salinos tienen en la estabilidad del homeodominio.
- Como paso final, las propiedades de unión a ADN de una selección de mutantes, y en algunos casos su dependencia con la temperatura, fueron analizadas mediante microcalorimetría (ITC).

3 *RESULTADOS*

3.1

ANÁLISIS TERMODINÁMICO Y ESTRUCTURAL DE LAS VARIANTES DEL HOMEODOMINIO DE HESX-1

3.1.1. Parámetros termodinámicos del homeodominio de HESX-1 salvaje

En la Figura 3.1 se muestran las curvas de desnaturalización térmica y química obtenidas para el homeodominio salvaje de HESX-1. A partir de las primeras se puede deducir una T_m de 56,3 °C a 0 M NaCl. Este valor concuerda con el descrito previamente por nuestro grupo de investigación (de la Mata *et al.*, 2002), y sitúa a HESX-1 entre los homeodominios más estables descritos hasta la fecha. Al aumentar la concentración de sal, se comprueba un claro aumento de estabilidad, en línea con lo observado en otros homeodominios (ver la introducción). Este comportamiento sugiere un balance electrostático desfavorable en HESX-1, motivado en buena medida por su carga neta de +5 a pH 6,0. Por último, la estabilidad medida a partir de la curva de desnaturalización con urea es $\Delta G_F = -4,43$ kcal/mol, muy en concordancia con la medida usando GdmCl ($-4,71$ kcal/mol), valores también más altos de lo habitual en homeodominios.

3.1.2. Análisis de secuencias de homeodominios y diseño de los mutantes

Con objeto de diseñar las variantes del homeodominio de HESX-1 que serán posteriormente producidas y analizadas, llevamos a cabo un análisis previo de secuencias presentes en homeodominios naturales. Actualmente, la mayor base de datos con información relativa a homeoproteínas es el “Homeodomain Resource” (Banerjee-Basu *et al.*, 2000, 2001, 2003). Contiene más de mil secuencias completas de proteínas procedentes de 31 especies distintas. El primer paso en nuestro estudio consistió en determinar la frecuencia de aparición de puentes salinos y pares hidrófobos en las posiciones 19-30, 31-42 y 17-52, a partir de la información almacenada en el “Homeodomain Resource”.

Con el objetivo de simplificar, nuestro análisis se restringió a un grupo reducido de 754 proteínas, que incluía únicamente secuencias de homeodominios con la longitud canónica de 60 residuos, sin inserciones ni deleciones. Los resultados obtenidos se representan esquemáticamente en la [Figura 3.2](#).

Se puede observar que las posiciones 19-30 están conectadas por interacciones de puente salino de ambas polaridades en el 52,5% de los casos (396 secuencias), mostrando una cierta preferencia por aquellas combinaciones que incluyen residuos de carga negativa y positiva en las posiciones 19 y 30, respectivamente (33,8% de los casos, [Figura 3.2a](#)). Por el contrario, residuos alifáticos en estas posiciones, tales como los presentes en HESX-1, son relativamente infrecuentes (2,5%).

Como hemos mencionado en la introducción, la estructura de rayos X del homeodominio de Engrailed (Fraenkel *et al.*, 1998; Stollar *et al.*, 2003) pone de

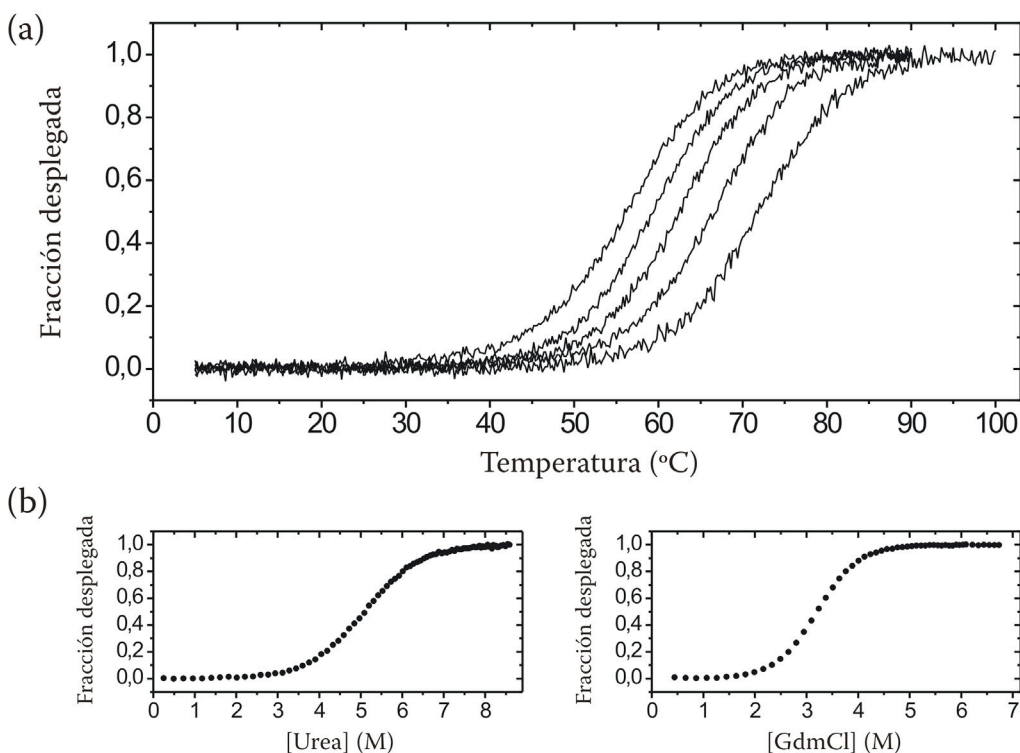


FIGURA 3.1: Curvas de desnaturalización correspondientes al homeodominio de HESX-1, obtenidas mediante medidas de dicroísmo circular. (a) Experimentos de desnaturalización térmica realizados en tampón fosfato 10 mM pH 6,0, 0,5 mM DTT, a concentraciones crecientes de NaCl: de izquierda a derecha, 0, 0,1, 0,4, 1 y 2 M NaCl. (b) Curvas de desnaturalización química con urea (panel izquierdo) o GdmCl (panel derecho), obtenidas a 20 °C en tampón fosfato 10 mM pH 6,0, 0,5 mM DTT, 100 mM NaCl.

manifiesto la existencia de una red cooperativa de puentes salinos que involucra no sólo las posiciones 19-30, sino también el par 15-37. De acuerdo con nuestro análisis, un patrón equivalente de interacciones que involucrase a la tétada 15-19-30-37 sería posible sólo en el 2,7% de las secuencias (ver [Figura 3.2b](#)). Curiosamente, una red alternativa que involucrase a la tríada 19-30-33 (normalmente E19-R30-E33) parece ser mucho más frecuente en esta familia

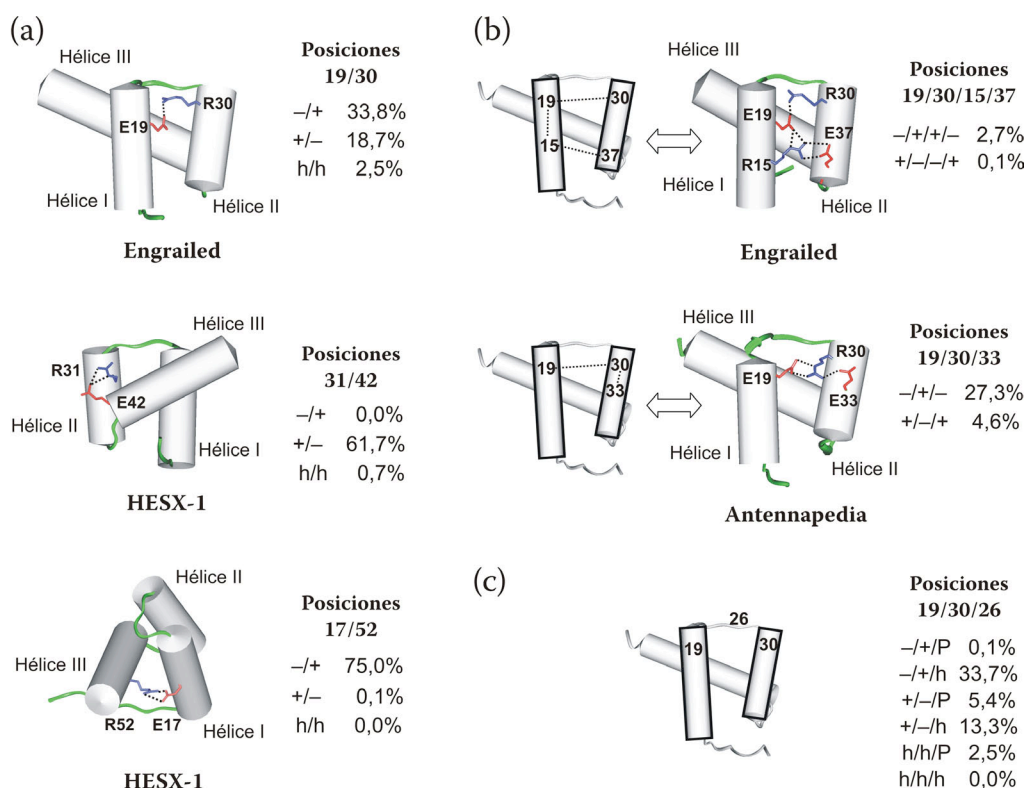


FIGURA 3.2: (a) Representación de las estructuras de los homeodominios de Engrailed (panel superior) y HESX-1 (paneles medio e inferior). Se destacan los puentes salinos interhélice más comúnmente observados en esta familia de proteínas. Dichas interacciones conectan las posiciones 19-30 (panel superior), 31-42 (panel medio) y 17-52 (panel inferior). A la derecha se muestran las frecuencias de aparición de puentes salinos respecto a pares hidrófobos en estas posiciones, de entre un grupo de 754 secuencias de homeodominios. (b) Representación esquemática de las redes de puentes salinos en que participan los residuos 19-30 más comunes en homeodominios. La frecuencia de aparición de cada patrón de interacciones en el conjunto de 754 secuencias se muestra a la derecha. (c) Correlación observada entre las posiciones 19-30 y el residuo 26 perteneciente al lazo I-II. Se muestra la frecuencia de aparición de cada combinación en el conjunto de 754 secuencias. Para (a), (b) y (c), los símbolos *h*, + y - se refieren a residuos alifáticos (V, L e I), cargados positivamente (R y K) y cargados negativamente (E y D), respectivamente. *P* significa prolina.

de proteínas, siendo posible en el 31,9% de las secuencias. Dicho puente salino bifurcado es claramente apreciable en la estructura del homeodominio de Antennapedia (ver Figuras 3.2b y 1.5b) (Billeter *et al.*, 1993; Fraenkel y Pabo, 1998).

Siguiendo con la interacción 19-30, su relación con el residuo 26 del lazo I-II ha sido objeto de cierto interés en el pasado. Mediante un análisis estadístico de 263 secuencias de homeodominios, Clarke identificó un patrón de covarianza centrado en dicho residuo (Clarke, 1995). De acuerdo con este análisis, los homeodominios fueron divididos en dos clases. La primera se caracteriza por presentar residuos alifáticos ramificados (L, V o I) en posición 26, mientras que la segunda incluye prolina en dicha posición. Un análisis posterior de 129 secuencias de homeodominios humanos representativos (Stollar *et al.*, 2003) reveló que el subgrupo con alifáticos ramificados en 26 normalmente incorporaba un puente salino conectando los residuos 19 y 30 (92% de los casos). Sorprendentemente, ninguno de los del subgrupo con prolina en 26, con más de 60 miembros, incluía esta potencial interacción. En base a esta observación, se sugirió un posible papel funcional para el par 19-30. En aquellos casos en que la posición 26 está ocupada por un residuo alifático, el puente salino 19-30 podría contribuir a la variabilidad conformacional del lazo I-II, afectando a la capacidad del homeodominio para llevar a cabo el ajuste inducido al unirse al ADN (Stollar *et al.*, 2003). Es de destacar que, hasta el momento de realizar nuestro trabajo, ninguna evidencia experimental, a favor o en contra de esta hipótesis, había sido presentada.

En relación a la correlación 19-26-30, nuestro análisis, que incluye un número mayor de secuencias (754), revela una situación algo más compleja (Figura 3.2c). Para homeodominios con un puente salino entre las posiciones 19 y 30, 26 puede ser tanto un aminoácido alifático ramificado (normalmente leucina) como prolina, dependiendo de la polaridad del puente salino. Si la carga positiva se encuentra en 30, 26 es un aminoácido alifático ramificado en más del 96% de los casos. Por el contrario, si la posición 30 tiene un residuo negativo, entonces 26 puede ser tanto leucina como prolina, si bien la leucina es claramente preferida. Por último, si los residuos 19 y 30 son alifáticos, 26 es siempre prolina (ver Figura 3.2c).

Por lo que respecta a 31 y 42, estas posiciones incorporan un puente salino en la gran mayoría de homeodominios (61,7%). En una pequeña fracción de secuencias, se puede encontrar un residuo alifático, ya sea en 31 (14,7%) o en 42 (3,8%). Por último, los dos son alifáticos en únicamente cinco secuencias dentro de nuestro conjunto de datos (esto es, el 0,7% de las secuencias analizadas) (Figura 3.2a).

De acuerdo con nuestro análisis, el puente salino más conservado en la familia de los homeodominios es aquel que conecta las posiciones 17 y 52, que

está presente en el 75,1% de las secuencias. Se pueden encontrar residuos hidrófobos en las posiciones 17 ó 52 en el 0,8% o en el 0,1% de los casos, respectivamente. No obstante, no aparece ningún par hidrófobo en estas posiciones en ninguna de las secuencias del conjunto que hemos analizado (Figura 3.2a).

En la Figura 3.3 se muestra una selección de estructuras representativas de homeodominios con puentes salinos 19-30, 31-42 y 17-52.

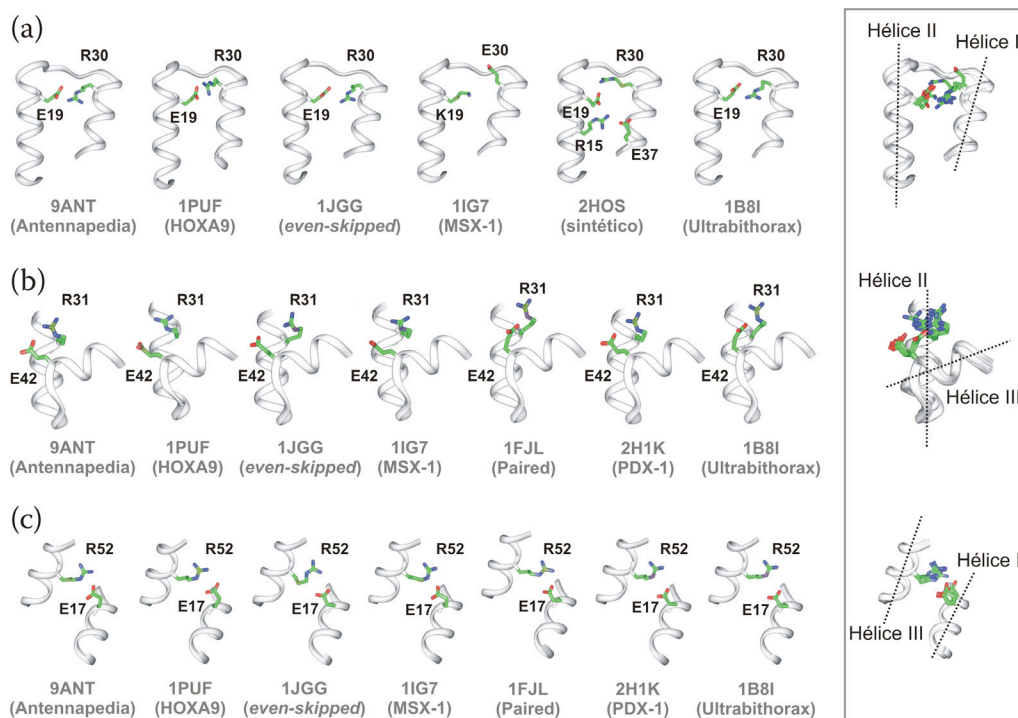


FIGURA 3.3: **Puentes salinos** entre las posiciones (a) 19-30, (b) 31-42 y (c) 17-52 observados en distintos homeodominios mediante cristalografía de rayos X (los códigos PDB aparecen indicados). En el panel de la derecha se muestran las superposiciones de las estructuras seleccionadas.

En base a este análisis de secuencias, decidimos sustituir los residuos presentes en las posiciones 15, 17, 19, 26, 30, 31, 37, 42 y 52 del homeodominio de HESX-1 salvaje por otros frecuentemente observados en homeodominios naturales, para dar 33 mutantes. Las 33 variantes producidas en este trabajo, representadas esquemáticamente en las Figuras 3.4 y 3.5 y numeradas del 1 al 33 como se muestra en la Tabla 3.1, se pueden englobar en cuatro grupos:

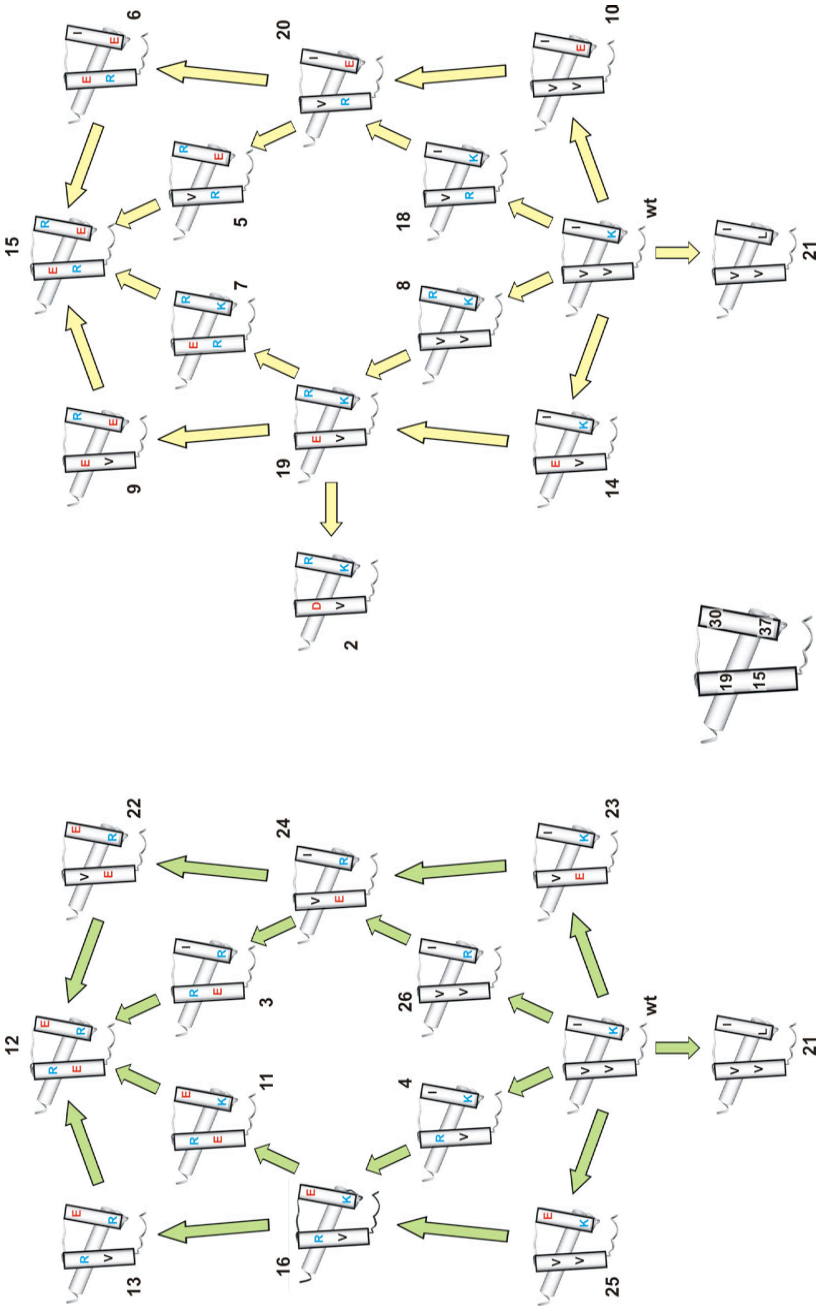


FIGURA 3.4: **Representación esquemática de las variantes del homeodominio de HESX-1** con sustituciones en las posiciones 19-30 y/o 15-37, producidas con objeto de satisfacer el objetivo principal, así como el punto 1 de los objetivos secundarios planteados en la sección 2.3. Se obtuvieron mutantes simples, dobles, triples y cuádruples. Los residuos cargados negativamente se muestran en rojo, los cargados positivamente en azul, y los alifáticos en negro. Panel izquierdo: variantes que conducen desde la proteína salvaje hasta el cuádruple mutante 12 (R19-E30, E15-R37). Panel derecho: variantes que conducen desde la proteína salvaje hasta el cuádruple mutante 15 (E19-R30, R15-E37). La variante 21 (mutante simple K37L) se muestra en ambos paneles.

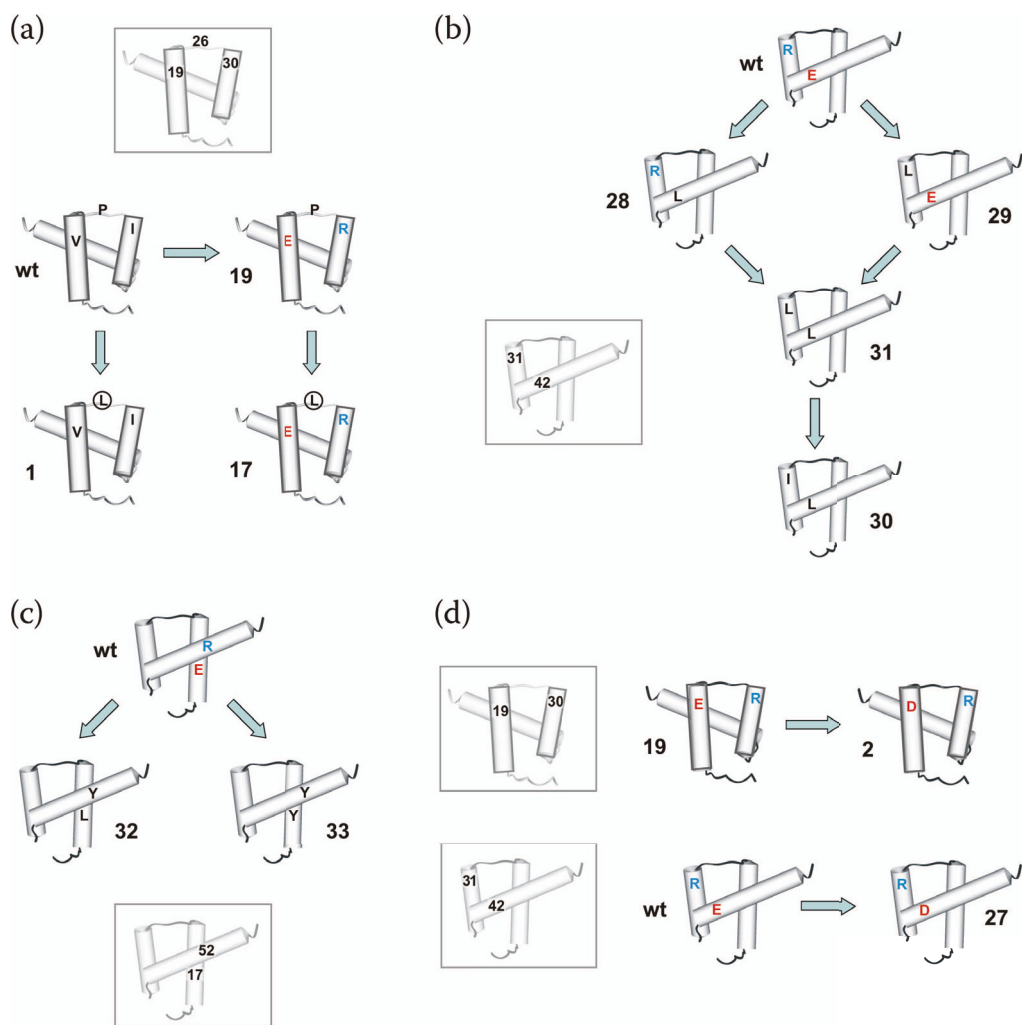


FIGURA 3.5: **Representación esquemática de las variantes del homeodominio de HESX-1** producidas con objeto de satisfacer los objetivos secundarios (a) 2, (b) 1 y 4, (c) 1, y (d) 3, planteados en el apartado 2.3.2. Se muestran las variantes con sustituciones en las posiciones (a) 19-26-30, (b) 31-42, (c) 17-52, y (d) con sustituciones E→D en los puentes salinos E19-R30 y R31-E42. Los residuos cargados negativamente se muestran en rojo, los cargados positivamente en azul, y los alifáticos en negro. [2]

- *Variantes con mutaciones en las posiciones 19-30 y 15-37:* Los pares V19-I30 y V15-K37 presentes en HESX-1, fueron secuencialmente sustituidos por diversos puentes salinos glutámico-arginina de distintas polaridades. El resultado fue la generación de un total de 25 variantes, incluyendo mutantes simples, dobles, triples y cuádruples (Figura 3.4). En casos favorables, la

inclusión de residuos cargados en las cuatro posiciones (19, 30, 15 y 37) simultáneamente podría permitir el establecimiento de redes cooperativas análogas a las observadas en Engrailed. Es de destacar que todas estas posiciones se localizan en la cara de HESX-1 opuesta a la superficie de interacción con el ADN, y por tanto deben modular la estabilidad del polipéptido sin afectar directamente sobre la estabilidad del complejo proteína/ADN. Por todo ello, los mutantes obtenidos tienen un claro interés en el estudio de la relación estabilidad-función planteado en la parte de objetivos. Además, nos dieron la oportunidad de conocer más acerca de la estabilización aportada por los puentes salinos en relación con los pares hidrófobos.

- *Variantes con mutaciones en las posiciones 19-26-30:* Mediante la sustitución P26L en la proteína salvaje y en la variante **19** (con el puente salino E19-R30), obtuvimos un total de cuatro combinaciones de residuos en estas posiciones, dos de ellas naturales y dos prácticamente inexistentes en homeodominios (Figura 3.5a). Su comparación en términos de estabilidad y afinidad por ADN nos acercó a entender las causas para la abundancia de una u otra combinación en la naturaleza.

- *Variantes con mutaciones en las posiciones 31-42 ó 17-52:* El puente salino R31-E42 fue sustituido por los pares alifáticos L31-L42 e I31-L42 (Figura 3.5b). Por su parte, el puente salino E17-R52 fue reemplazado por los pares Y17-Y52 y L17-Y52 (Figura 3.5c). En estos casos, el propósito volvió a ser comparar la eficacia estabilizante de interacciones electrostáticas e hidrófobas, pero centrándonos esta vez en puentes salinos más conservados.

- *Variantes con mutaciones E→D en los puentes salinos E19-R30 y R31-E42:* Las dos nuevas proteínas generadas con estas mutaciones (Figura 3.5d) nos permitieron comparar puentes salinos diferenciados por un único grupo metileno.

3.1.3. Análisis de la estabilidad de las variantes de HESX-1: aspectos metodológicos

Con objeto de llevar a cabo un análisis de estabilidad detallado, las 33 variantes producidas fueron sometidas a experimentos de desnaturalización tanto térmica como química. En todos los casos, el desplegamiento de los diferentes homeodominios se monitorizó empleando dicroísmo circular. Para cada mutante, las curvas de desnaturalización térmica se adquirieron a cinco concentraciones de cloruro sódico distintas, en el rango de 0 a 2 M (0, 0,1, 0,4, 1,0 y 2,0 M). A partir de estas curvas, derivamos los valores de ΔG_F como se explica en la parte experimental, y analizamos su dependencia con la fuerza iónica a una temperatura común de 55 ó 77 °C, según el grupo de mutantes (ver

Tabla 3.1). El empleo de la misma temperatura para cada conjunto de mutantes facilitó el análisis comparativo de los mismos. El efecto de la fuerza iónica sobre ΔG_F se tomó como una medida de la contribución de las interacciones electrostáticas al plegamiento de los diferentes homeodominios (ver parte experimental). A modo de ejemplo, en las Figuras 3.1 y 3.6 se muestran las variaciones en la elipticidad a 222 nm observadas para los polipéptidos salvaje y las variantes 24, 22, 3 y 12, en función de la temperatura o la concentración de agente desnaturante. Los parámetros termodinámicos derivados del análisis de estos experimentos se muestran en la **Tabla 3.1**. En ella, las 33 variantes analizadas se presentan divididas en grupos dependiendo de las posiciones

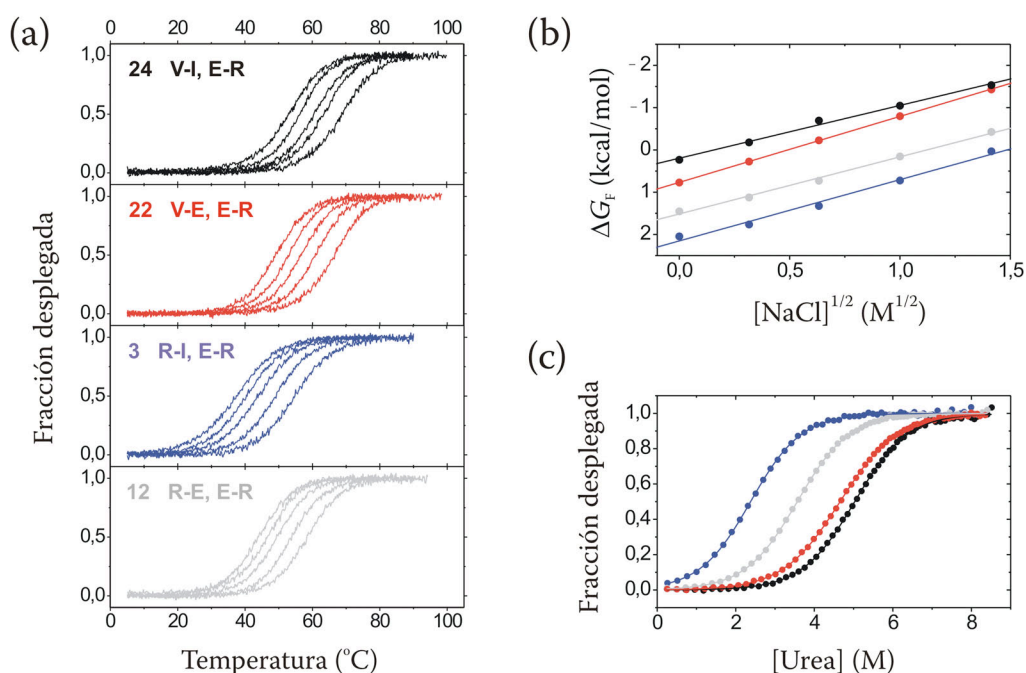


FIGURA 3.6: Análisis termodinámico de las variantes del homeodominio de HESX-1 24 (negro), 22 (rojo), 3 (azul) y 12 (gris). (a) Curvas de desnaturalización térmica obtenidas mediante dicroísmo circular a diferentes concentraciones de NaCl (0, 0,1, 0,4, 1,0 y 2,0 M), en tampón fosfato 10 mM pH 6,0, 0,5 mM DTT. Puede apreciarse como, en todos los casos, la sal aumenta la estabilidad térmica. (b) A partir de las curvas de desnaturalización térmica se derivaron los valores de ΔG_F a una temperatura común (55 °C), tal y como se describe en la parte experimental. En todos los casos, ΔG_F exhibe una dependencia lineal con la raíz cuadrada de la fuerza iónica, lo que sugiere que el apantallado electrostático de interacciones desfavorables es el mecanismo dominante de la estabilización por sal observada. (c) Curvas de desnaturalización química con urea medidas mediante dicroísmo circular a 20 °C, en tampón fosfato 10 mM pH 6,0, 0,5 mM DTT, 100 mM NaCl.

TABLA 3.1: Datos de estabilidad medidos para HESX-1 salvaje y las variantes **1** a **33** a partir de experimentos de desnaturalización química (izquierda, empleando urea o GdmCl según se indica) y térmica (derecha)^a

		Desnaturalización química		Desnaturalización térmica				
Variante	Residuos 19-30, 15-37	ΔG_F^b (kcal/mol) Urea	m^c	T_m^d (°C) 0/2 M NaCl	ΔG_F^e (kcal/mol) a 55 °C			m^f
					0 M NaCl total	2 M NaCl no polar	Coulómbico	
1	V-I,V-K ^(L26)	-	-	-	-	-	-	-
2	D-R, V-K	-2,01	0,8	36,2/53,6	2,53±0,27	0,19±0,04	2,34±0,27	-1,7
3	R-I, E-R	-2,14	0,9	37,5/54,7	2,04±0,24	0,03±0,03	2,01±0,24	-1,5
4	R-I, V-K	-2,01	0,9	37,7/58,4	2,14±0,24	-0,39±0,04	2,53±0,24	-1,9
5	V-R, R-E	-2,52	0,8	38,3/62,8	2,18±0,23	-1,00±0,09	3,18±0,25	-2,3
6	E-I, R-E	-3,18	0,8	41,0/60,8	1,77±0,18	-0,76±0,07	2,53±0,19	-1,9
7	E-R, R-K	-2,54	0,9	41,2/58,6	1,89±0,18	-0,50±0,05	2,39±0,19	-1,7
8	V-R, V-K	-2,59	0,8	41,5/64,3	1,50±0,17	-1,01±0,10	2,51±0,20	-1,8
9	E-R, V-E	-3,39	0,8	41,8/64,8	2,18±0,18	-1,44±0,11	3,62±0,21	-2,6
10	V-I, V-E	-3,84	0,8	43,7/69,9	1,93±0,15	-1,63±0,17	3,56±0,23	-2,7
11	R-E, E-K	-3,12	0,9	43,7/58,7	1,50±0,14	-0,46±0,05	1,96±0,15	-1,4
12	R-E, E-R	-3,14	0,9	44,0/58,5	1,45±0,14	-0,43±0,05	1,88±0,15	-1,4
13	R-E, V-R	-3,23	0,9	44,3/62,2	1,39±0,14	-0,92±0,08	2,31±0,16	-1,7
14	E-I, V-K	-3,23	0,8	44,8/63,5	1,33±0,13	-1,05±0,09	2,38±0,16	-1,7
15	E-R, R-E	-3,57	0,9	45,2/62,8	1,45±0,13	-1,06±0,09	2,51±0,16	-1,8
16	R-E, V-K	-3,32	0,9	46,0/63,1	1,28±0,12	-1,09±0,09	2,37±0,15	-1,7
17	E-R, V-K ^(L26)	-3,53	0,8	46,3/65,0	1,11±0,11	-1,80±0,12	2,91±0,16	-2,1
18	V-I, R-K	-3,56	1,0	47,1/64,3	1,02±0,10	-1,05±0,10	2,07±0,14	-1,5
19	E-R, V-K	-3,69	0,9	48,4/66,6	0,89±0,09	-1,50±0,13	2,39±0,16	-1,7
20	V-I, R-E	-3,80	0,8	48,6/67,9	0,88±0,09	-1,51±0,15	2,39±0,17	-1,7
21	V-I, V-L	-4,60	0,8	48,9/70,8	0,64±0,07	-1,46±0,18	2,10±0,19	-1,5
22	V-E, E-R	-3,75	0,8	49,2/66,6	0,77±0,08	-1,43±0,13	2,20±0,15	-1,6
23	V-I, E-K	-4,17	0,9	52,3/68,8	0,33±0,05	-1,70±0,16	2,03±0,17	-1,4
24	V-I, E-R	-4,38	0,9	53,1/68,9	0,23±0,04	-1,53±0,16	1,76±0,17	-1,3
25	V-E, V-K	-4,20	0,9	53,2/70,2	0,23±0,04	-1,83±0,17	2,06±0,17	-1,5
26	V-I, V-R	-5,11	1,0	56,1/72,9	-0,13±0,03	-2,05±0,22	1,92±0,22	-1,4
wt	V-I, V-K	-4,43	0,9	56,3/72,0	-0,16±0,03	-1,94±0,20	1,78±0,20	-1,3
Variante	Residuos 31-42	ΔG_F^b (kcal/mol) Urea	m^c	T_m^d (°C) 0/2 M NaCl	ΔG_F^e (kcal/mol) a 55 °C			m^f
					0 M NaCl total	2 M NaCl no polar	Coulómbico	
27	R-D	-3,70	0,9	50,1/66,1	0,56±0,06	-1,28±0,12	1,84±0,13	-1,3
wt	R-E	-4,43	0,9	56,3/72,0	-0,16±0,03	-1,94±0,20	1,78±0,20	-1,3
Variante	Residuos 31-42	ΔG_F^b (kcal/mol) GdmCl	m^c	T_m^d (°C) 0/2 M NaCl	ΔG_F^e (kcal/mol) a 77 °C			m^f
					0 M NaCl total	2 M NaCl no polar	Coulómbico	
wt	R-E	-4,71	1,5	56,3/72,0	3,00±0,29	0,68±0,07	2,32±0,30	-1,7
28	R-L	-5,03	1,4	58,2/75,8	2,57±0,25	0,21±0,04	2,36±0,25	-1,7
29	L-E	-6,52	1,5	71,7/85,2	0,68±0,07	-1,04±0,09	1,72±0,11	-1,3
30	I-L	-8,44	1,7	82,8/91,6	-0,75±0,07	-1,91±0,16	1,16±0,17	-0,8
31	L-L	-9,48	1,6	88,7/95,9	-1,26±0,12	-2,43±0,22	1,17±0,25	-0,8
Variante	Residuos 17-52	ΔG_F^b (kcal/mol) Urea	m^c	T_m^d (°C) 0 M NaCl	ΔG_F^e (kcal/mol) a 55 °C			m^f
					0 M NaCl total	2 M NaCl no polar	Coulómbico	
wt	E-R	-4,43	0,9	56,3	-0,16±0,03	-1,94±0,20	1,78±0,20	-1,3
32	L-Y	-	-	-	-	-	-	-
33	Y-Y	-	-	41,3	-	-	-	-

^{a, b, c, d, e, f} (PÁGINA SIGUIENTE)

afectadas por las mutaciones, y ordenadas dentro de cada uno por T_m creciente a 0 M NaCl.

Estudios recientes han mostrado que interacciones electrostáticas favorables presentes en el estado desplegado pueden tener una influencia significativa en la estabilidad de la proteína (Pace *et al.*, 2000). En el caso de los mutantes de HESX-1, la constancia en los valores m que hemos medido por experimentos de desnaturalización química (m es la pendiente de la recta de la dependencia de ΔG_F con la concentración de agente desnaturalizante) se puede tomar como evidencia contra esta posibilidad, y sugiere que las diferencias en estabilidad observadas reflejan principalmente las interacciones establecidas por los residuos mutados en el estado nativo.

Por lo que respecta a los experimentos de desnaturalización térmica, en todos los casos se aprecia una acusada estabilización de los polipéptidos al incrementar la concentración de NaCl (tal y como se había observado previamente para la proteína salvaje [Figura 3.1]). Además, los valores de ΔG_F muestran una dependencia aproximadamente lineal con la raíz cuadrada de la fuerza iónica, como se puede comprobar en la Figura 3.6b. Esta observación sugiere que el mecanismo dominante de la estabilización por sal obtenida es el apantallado electrostático de interacciones desfavorables (Debye y Huckel, 1923; Gloss y Placek, 2002; de los Rios y Plaxco, 2005).

TABLA 3.1 (CONT.)

^a Todos los experimentos se realizaron en tampón fosfato 10 mM pH 6,0, 0,5 mM DTT, a la concentración de NaCl indicada (100 mM en el caso de los experimentos de desnaturalización química). Los mutantes están numerados de acuerdo con su estabilidad térmica a 0 M NaCl. Todas las variantes tienen prolina en la posición 26 del lazo I-II, a menos que se indique expresamente (entre paréntesis para las variantes 1 y 17).

^b ΔG_F es la variación de energía libre de Gibbs de plegamiento a 20 °C. Sus errores, derivados de los ajustes de las curvas, fueron siempre menores de 0,15 kcal/mol.

^c m (kcal/[mol·M]) es la pendiente de la recta que describe la dependencia entre ΔG_F y la concentración de agente desnaturalizante. Sus errores, derivados de los ajustes de las curvas, fueron siempre inferiores a 0,03 kcal/(mol·M).

^d T_m es el punto medio de la transición de desnaturalización térmica. Se muestran los valores medidos a 0 y a 2 M NaCl, separados por barras. Sus errores, derivados de los ajustes de las curvas, fueron siempre menores de 0,1 °C.

^e ΔG_F es la variación de energía libre de Gibbs de plegamiento a 55 °C ó 77 °C (según se indica). De acuerdo con la aproximación empleada, ΔG_F a 0 M NaCl es la variación total, ΔG_F a 2 M NaCl representa la contribución no polar, y la diferencia entre ambos valores representa la contribución coulombica a ΔG_F .

^f $m1$ (kcal/[mol·M^{1/2}]) es la pendiente de la recta que describe la dependencia entre ΔG_F y $[\text{NaCl}]^{1/2}$, determinada como describen de los Rios y Plaxco (de los Rios y Plaxco, 2005). Este parámetro es proporcional al término $\Delta G_{\text{coulomb}}$.

Siguiendo la metodología descrita por Schmid y colaboradores (Perl *et al.*, 2000; Martin *et al.*, 2002; Wunderlich *et al.*, 2005; Max *et al.*, 2007), asumimos que en el ΔG_F medido a 0 M NaCl están presentes los aportes procedentes de interacciones tanto electrostáticas como no polares. A este término lo llamaremos ΔG_{total} . Las contribuciones de las mutaciones a la estabilidad que permanecen a 2 M NaCl reflejan primariamente las diferencias en empaquetamiento y en interacciones hidrófobas, a lo que llamaremos $\Delta G_{\text{no polar}}$. Otros efectos, tales como los costes de desolvatación, las propensiones de hélice intrínsecas de los diferentes residuos o sus entropías conformacionales, se pueden incluir también en este término. Finalmente, aquellas contribuciones que resultan apantalladas a 2 M NaCl se consideraron interacciones coulómbicas. A este término lo llamaremos $\Delta G_{\text{coulomb}}$:

$$\begin{aligned}\Delta G_{\text{total}} &= \Delta G_{F \text{ 0 M NaCl}} \\ \Delta G_{\text{no polar}} &= \Delta G_{F \text{ 2 M NaCl}} \\ \Delta G_{\text{coulomb}} &= \Delta G_{F \text{ 0 M NaCl}} - \Delta G_{F \text{ 2 M NaCl}}\end{aligned}$$

A lo largo de la presentación de resultados, los residuos de cada variante relevantes para la discusión se indican entre paréntesis. Para los valores de ΔG destacados en el texto principal, los errores han sido omitidos para facilitar la lectura. Éstos pueden consultarse en la [Tabla 3.1](#).

3.1.4. Puentes salinos frente a pares hidrófobos en las posiciones 19-30

Se analizaron 25 variantes del homeodominio de HESX-1 con diferentes combinaciones de residuos cargados o alifáticos en las posiciones 19-30 y 15-37. Dichos polipéptidos se representan de forma esquemática en la [Figura 3.4](#). De ellos, 14 incluyen potenciales puentes salinos de diferentes polaridades, conectando las posiciones 19-30 (variantes **2**, **7**, **9**, **13**, **16** y **19**) ó 15-37 (variantes **3**, **5**, **6**, **11**, **20**, **22**, **23** y **24**). Es de destacar que la presencia de puentes salinos de polaridad opuesta conectando los pares 19-30 y 15-37 simultáneamente (en los mutantes **11**, **12** y **15**) permitiría, en principio, el establecimiento de una red cooperativa análoga a la anteriormente descrita para el homeodominio de Engrailed ([Figura 1.5](#) y [Figura 3.2b](#)), y presente en una fracción mínima de homeodominios naturales de acuerdo con nuestra inspección del “Homeodomain Resource”. Está bien establecido que, en casos favorables, la formación de redes cooperativas puede suponer una considerable mejora sobre el efecto estabilizante de los puentes salinos individuales (Horovitz *et al.*, 1990; Spek *et al.*, 1998; Kumar y Nussinov, 1999; Karshikoff y

Ladenstein, 2001). En principio, las diferentes combinaciones de residuos alifáticos y cargados que presentan nuestras variantes, nos deben permitir ensayar la influencia de puentes salinos (de diferentes polaridades, bien aislados o formando parte de redes cooperativas) y pares hidrófobos en estas cuatro posiciones sobre la estabilidad de HESX-1. A su vez, las mutaciones seleccionadas nos deben permitir generar polipéptidos de estabilidades diversas sin afectar de modo directo a sus propiedades de unión a ADN, ya que, tal y como hemos mencionado anteriormente, las mutaciones involucran posiciones alejadas de la zona de reconocimiento.

Los resultados obtenidos se representan en la [Tabla 3.1](#) y se resumen en las [Figuras 3.7](#) y [3.8](#). Se puede observar que las sustituciones simples o dobles de los residuos V19-I30 por cadenas laterales cargadas, para construir puentes salinos de diferentes polaridades resultan, en todos los casos, en una clara desestabilización del plegamiento del homeodominio. Curiosamente, este efecto parece relativamente independiente del contexto aportado por el par 15-37. La dependencia entre la fuerza iónica y ΔG_F a 55 °C nos proporciona pistas sobre el origen de este efecto. Así, pese a que la influencia desfavorable de las mutaciones queda, en la mayoría de los casos, ligeramente atenuada a 2M NaCl (lo que indica un balance electrostático menos optimizado en los mutantes), el efecto desestabilizante persiste en gran medida a esa concentración de sal (ver los términos $\Delta G_{\text{coulomb}}$ y $\Delta G_{\text{no polar}}$ en la [Tabla 3.1](#)). De hecho, podemos observar en la [Figura 3.7](#) que los puntos correspondientes al término $\Delta G_{\text{no polar}}$ siguen, por lo general, el mismo patrón que los correspondientes a ΔG_{total} . En cambio, los puntos correspondientes a $\Delta G_{\text{coulomb}}$ sólo exhiben pequeñas fluctuaciones, con valores de $\Delta \Delta G_{\text{coulomb}}$ próximos a cero. Asimismo, en la [Figura 3.8](#) se aprecia una correlación claramente más alta entre ΔG_{total} y $\Delta G_{\text{no polar}}$ que entre ΔG_{total} y $\Delta G_{\text{coulomb}}$. En conjunto, estas observaciones parecen sugerir que las causas principales de la desestabilización motivada por las mutaciones son de origen no electrostático.

Una explicación razonable para este resultado es que V19 e I30 establecen unas interacciones de empaquetamiento con el núcleo hidrófobo de la proteína más optimizadas que E o R. El mejor empaquetamiento conseguido por V19 e I30 es suficiente para compensar sus menores propensiones intrínsecas de hélice (tanto V como I están ramificados en β y menos favorecidos para formar hélice α que R o E [O'Neil y DeGrado, 1990; Chakrabartty *et al.*, 1994; Pace y Scholtz, 1998]). Incluso las variantes **11** (R19/E30/E15/K37), **12** (R19/E30/E15/R37) y **15** (E19/R30/R15/E37), en las que es posible el establecimiento de una red de puentes salinos entre las posiciones 15, 19, 30 y 37, se ven severamente desestabilizadas con respecto al polipéptido salvaje. De acuerdo con los datos obtenidos, la presencia simultánea de puentes salinos de polaridad opuesta en las posiciones 19-30 y 15-37 no se traduce en

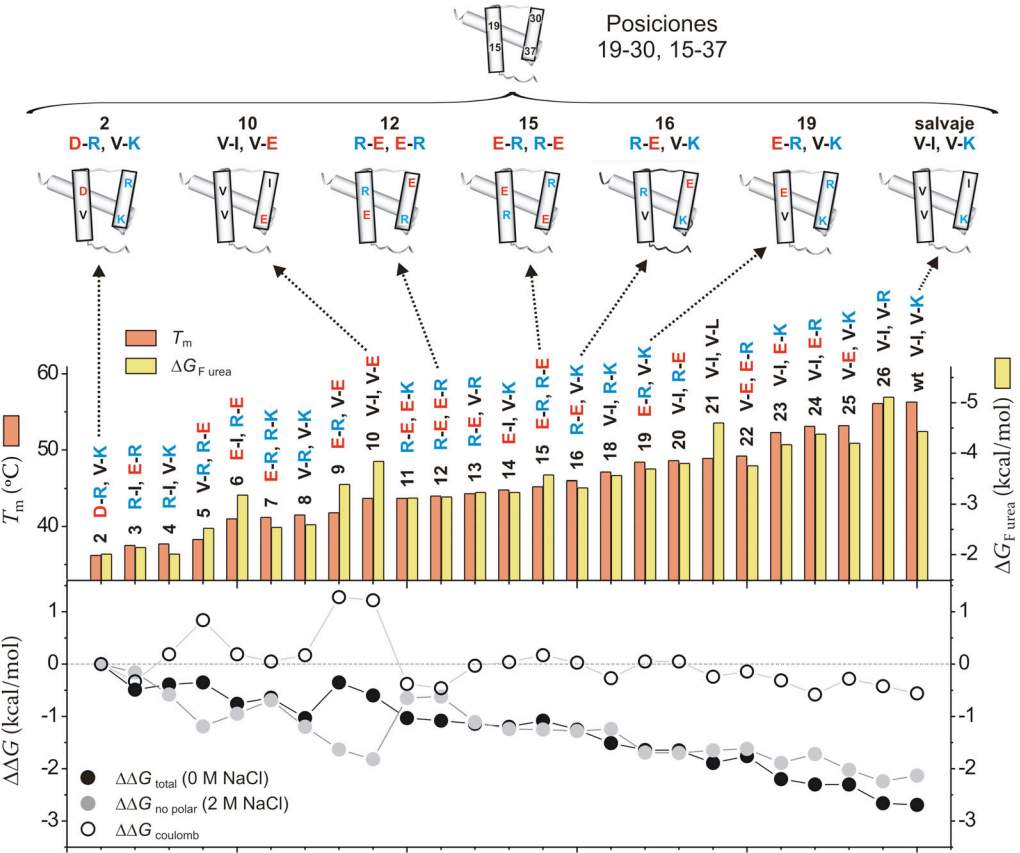


FIGURA 3.7: Representación gráfica de los parámetros termodinámicos medidos para las variantes con sustituciones en las posiciones 19-30, 15-37. En el gráfico de columnas se representan, en rojo, las T_m medidas a 0 M NaCl en tampón fosfato 10 mM pH 6,0, 0,5 mM DTT, y en amarillo, los ΔG_F medidos mediante desnaturalización con urea a 20 °C, en tampón fosfato 10 mM pH 6,0, 0,5 mM DTT, 100 mM NaCl. En el panel inferior se representan las diferencias de energía libre a 55 °C respecto a la variante menos estable (variante 2), derivadas de los experimentos de desnaturalización térmica a 0 M NaCl ($\Delta\Delta G_{total}$), a 2 M NaCl ($\Delta\Delta G_{no\ polar}$) y la diferencia entre ambos ($\Delta\Delta G_{coulomb}$). Valores negativos indican una energía libre más favorable que en la variante 2. Los errores no se representan para favorecer la claridad. En la parte superior se destacan esquemáticamente algunas de las variantes. Los residuos cargados negativamente se muestran en rojo, los cargados positivamente en azul, y los alifáticos en negro.

una mejora evidente de los términos $\Delta G_{coulomb}$ para estas variantes, que, en cambio, vuelven a presentar contribuciones $\Delta G_{no\ polar}$ claramente desfavorables.

Por lo que respecta a la polaridad de los puentes salinos introducidos en HESX-1, los resultados obtenidos muestran, además, cierta preferencia por la combinación E19/R30 respecto a la combinación R19/E30. De hecho, las

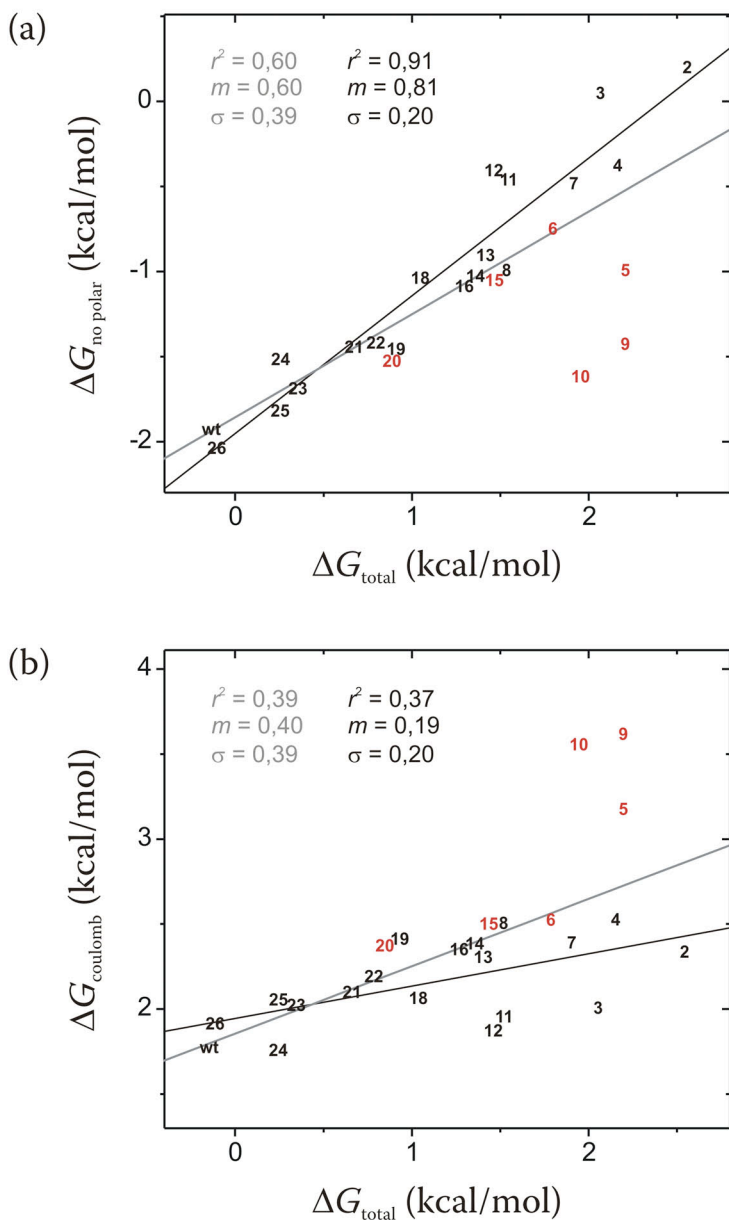


FIGURA 3.8: **Gráficas de correlación** entre ΔG_{total} y los términos (a) $\Delta G_{\text{no polar}}$ y (b) $\Delta G_{\text{coulomb}}$, a 55 °C, obtenidas para las variantes salvaje, 2 a 16 y 18 a 26 (con sustituciones en las posiciones 19-30, 15-37). Los números en el interior de las gráficas identifican las variantes. En rojo se muestran las que llevan ácido glutámico en posición 37. En las dos gráficas y en ambos ejes se ha usado la misma escala para facilitar las comparaciones. Las líneas representan los ajustes de los datos por mínimos cuadrados, considerando el conjunto completo de datos (gris), y un conjunto reducido que excluye las variantes representadas en rojo (negro). r^2 , coeficiente de correlación elevado al cuadrado. m , pendiente de la recta. σ , desviación típica.

variantes **19** (E19/R30/V15/K37) y **15** (E19/R30/R15/E37) son más estables que la **16** (R19/E30/V15/K37) y la **12** (R19/E30/E15/R37), respectivamente (ver [Tabla 3.1](#)). La diferencia entre ambas combinaciones es, nuevamente, de origen primordialmente no polar. Es probable que la distribución concreta de los residuos en las posiciones 19 y 30 presente en **19** permita unas mejores interacciones de empaquetamiento con el núcleo hidrófobo de la proteína, y/o una interacción más optimizada entre la posición 30 y el lazo I-II. Es interesante comprobar que esta preferencia se ve reflejada en la abundancia relativa de ambos tipos de puentes salinos en la base de datos “Homeodomain Resource”. De esta forma, y como se ha indicado previamente, puentes salinos de polaridad equivalente a la presente en el mutante **19** (residuo positivo en 30 y negativo en 19) y a la presente en el mutante **16** (residuo positivo en 19 y negativo en 30) están presentes en el 33,8% y en el 18,7% de las secuencias, respectivamente (ver [Figura 3.2a](#)).

Por último, en las [Figuras 3.7](#) y [3.8](#) podemos detectar algunas variantes que se separan de la tendencia general descrita, en la cual el término $\Delta G_{\text{no polar}}$ domina las diferencias de estabilidad observadas. Dichas variantes (en rojo en la [Figura 3.8](#)) muestran como rasgo común la presencia de ácido glutámico en posición 37. En ellas, la contribución $\Delta G_{\text{coulomb}}$ es especialmente desfavorable, y constituye el principal motivo de su baja estabilidad. La explicación podría encontrarse en que la introducción de una cadena de ácido glutámico en la posición 37 suprime el contacto intrahélice D33-K37, presente en el polipéptido natural (ver [Figura 3.23](#), más adelante), sustituyéndola por una interacción repulsiva D33-E37. Este efecto parece menos acusado en los casos en que E37 aparece acompañado de una arginina en posición 15, presumiblemente porque el puente salino R15-E37 podría suponer una compensación electrostática a la repulsión E37-D33.

En conclusión, interacciones hidrófobas sencillas establecidas por residuos alifáticos en las posiciones 19 y 30 posibilitan una estabilización del plegamiento de HESX-1 más efectiva que puentes salinos de cualquier polaridad en dichas posiciones, ya sean aislados o formando redes.

3.1.5. Interacción entre el par 19-30 y el residuo 26 del lazo I-II: influencia en la estabilidad del homeodominio

Como se ha mencionado previamente, una inspección cuidadosa de la base de datos “Homeodomain Resource” revela una clara e intrigante correlación entre los residuos 19-30 y el residuo 26, localizado en el lazo I-II. Además del papel propuesto en el ajuste inducido que hemos comentado (*Stollar et al.*, 2003), la red de interacciones centrada en estos tres residuos, así como el empa-

quetamiento entre sus cadenas laterales, podrían tener también un efecto significativo en la estabilidad de la proteína.

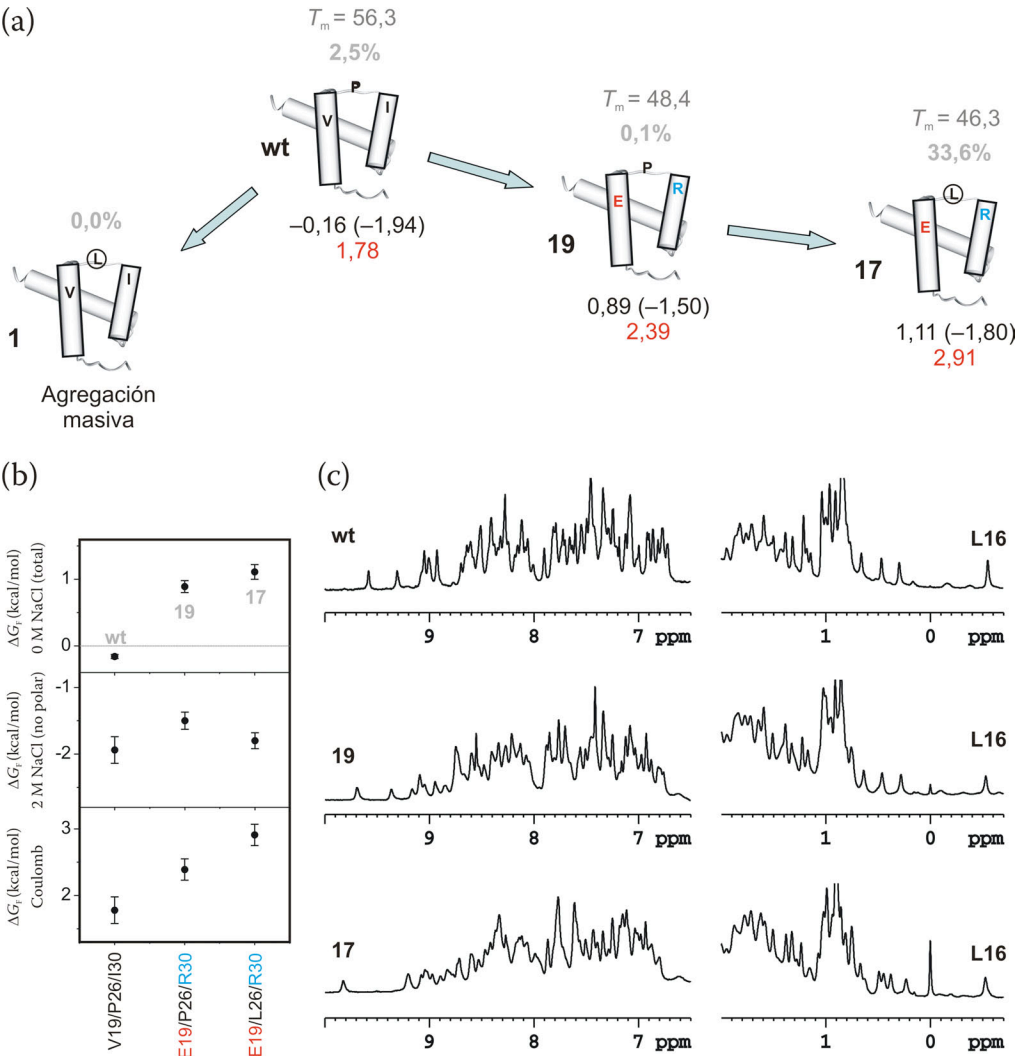
El homeodominio de HESX-1 salvaje incluye una prolina en posición 26 y los residuos 19 y 30 son alifáticos (V19-I30, ver [Figura 3.5a](#)). Esta pareja fue sustituida por un puente salino E19-R30 en el mutante **19**, previamente analizado, lo que resultó en una combinación de residuos inusual (E19/P26/R30). Con objeto de determinar el efecto de puentes alifáticos en relación a puentes salinos conectando las posiciones 19-30 en la estabilidad proteica, en el contexto de prolina o leucina en posición 26, dos mutantes adicionales fueron producidos y analizados ([Figura 3.5a](#)). En primer lugar, obtuvimos el triple mutante **17**, que contiene la tríada más frecuentemente encontrada en homeodominios (E19/L26/R30). En segundo lugar, P26 fue sustituida por leucina en la proteína salvaje para producir el mutante **1**, caracterizado por la combinación V19/L26/I30, no encontrada hasta la fecha en homeodominios naturales.

Como análisis preliminar, y con objeto de verificar que las mutaciones introducidas no tienen un efecto significativo sobre el plegamiento de la proteína, registramos espectros de RMN de 1D y de 2D para los mutantes **17** (E19/L26/R30) y **19** (E19/P26/R30) y los comparamos con los previamente obtenidos para el polipéptido salvaje. Los resultados obtenidos se muestran en la [Figura 3.9c](#). Puede apreciarse como todos los espectros presentan señales finas con una dispersión similar, lo que descarta la existencia de agregación o de heterogeneidad conformacional para las tres proteínas. Además, señales distinguidas por desplazamientos químicos peculiares (como los de los grupos metilo de L16), características de la estructura del homeodominio, pudieron ser fácilmente identificadas en todos los casos. Asimismo, ciertos NOEs clave de larga distancia, que son definitorios de esta familia de proteínas, pudieron ser asignados también para los dos mutantes, lo que descarta la existencia de diferencias estructurales pronunciadas entre ellos y el polipéptido salvaje (resultados no mostrados).

En una siguiente etapa, los mutantes **17** y **19** fueron sujetos a estudios de desnaturalización. Los resultados obtenidos se muestran esquemáticamente en la [Figura 3.9, a y b](#). Se puede observar que la sustitución P26L en **19** (E19/P26/R30), para dar **17** (E19/L26/R30), tiene un efecto menor en la estabilidad del homeodominio. No obstante, pese al carácter conservador de dicha mutación, el polipéptido **17** es ligeramente menos estable que **19**, de acuerdo con experimentos de desnaturalización tanto química como térmica. El comportamiento de ΔG_F a 55 °C en función de la presencia de sal sugiere un balance electrostático menos favorable en **17** como el principal motivo de esta leve desestabilización (ver los términos $\Delta G_{\text{coulomb}}$ en la [Tabla 3.1](#) y en la [Figura 3.9, a y b](#)). Una explicación especulativa pero plausible a este resultado es que la

introducción de leucina en lugar de P26 induce pequeñas diferencias conformacionales en el lazo I-II, que podrían conducir a una interacción electrostática menos eficaz entre E19 y R30. Adicionalmente, conviene tener en cuenta que la mayor estabilidad de la variante **19** puede deberse en parte a la menor libertad conformacional de la prolina respecto a la leucina, que supone una menor pérdida de entropía asociada al plegamiento.

Según nuestros resultados, pese al abrumador predominio de la tríada E19/L26/R30 respecto a E19/P26/R30 en secuencias naturales, ésta primera no parece tener un efecto beneficioso sobre la estabilidad conformacional del polipéptido (al menos en el caso de HESX-1). Esta observación refuerza la



hipótesis propuesta por Luisi y colaboradores, de acuerdo con la cual la elevada frecuencia de aparición de la combinación E19/L26/R30 en homeodominios sería el reflejo de su papel funcional en el reconocimiento de ADN (Stollar *et al.*, 2003).

Finalmente, en el caso de la variante 1 (V19/L26/I30) se observó un comportamiento muy diferente. Sorprendentemente, la sustitución P26L sobre la proteína salvaje indujo la formación de cuerpos de inclusión bajo todas las condiciones de expresión ensayadas. Además, la pequeña fracción de proteína expresada en forma soluble mostró una fuerte tendencia a agregar y precipitar durante el proceso de purificación, siendo imposible su aislamiento en todos los intentos llevados a cabo. Por lo tanto, este mutante no pudo ser sujeto a un análisis termodinámico detallado. Sin embargo, teniendo en cuenta que el residuo 26 no está expuesto al disolvente en la proteína nativa, y que por tanto no es probable que afecte significativamente a la solubilidad de la misma, este comportamiento sugiere que la mutación P26L desestabiliza profundamente el homeodominio.

En conclusión, una prolina en posición 26 es compatible tanto con residuos alifáticos como polares (neutros o cargados) en 19-30. Por el contrario, la cadena lateral de leucina en 26, más grande y conformacionalmente más móvil que la prolina, es compatible únicamente con residuos polares en las posiciones 19-30. Violaciones de esta regla conducen a una severa desestabilización y/o agregación de la proteína. Este comportamiento es consistente con la información almacenada en el “Homeodomain Resource” y explica por

FIGURA 3.9 (PÁGINA ANTERIOR): Influencia del par 19-30 sobre la estabilidad de HESX-1 en el contexto de prolina o leucina en la posición 26. (a) Representación esquemática del homeodominio de HESX-1 salvaje (wt) junto con las variantes 1, 19 y 17, caracterizadas por diferentes combinaciones “naturales” (como en la proteína salvaje y en la variante 17) y “no naturales” (como en las variantes 1 y 19) de los residuos en posiciones 19, 30 y 26. Los residuos cargados negativamente se muestran en rojo, y los cargados positivamente en azul. El porcentaje en que se dan combinaciones del mismo tipo en homeodominios, de acuerdo con la Figura 3.2, se representa sobre cada variante en gris claro. Asimismo, las temperaturas de desnaturalización (T_m) a 0 M NaCl se indican en gris. Bajo cada variante, se dan los valores numéricos de ΔG_{total} , $\Delta G_{no\ polar}$ (entre paréntesis) y $\Delta G_{coulomb}$ (en rojo) a 55 °C, obtenidos a partir de experimentos de desnaturalización térmica. Para la variante 1, no se pudo analizar la estabilidad debido a la pronunciada tendencia a agregar y precipitar de la proteína. Los detalles experimentales se recogen en la Tabla 3.1 y en la Figura 3.6. Las energías libres se dan en kcal/mol. (b) Representación gráfica de los valores de ΔG_F a 55 °C medidos a partir de experimentos de desnaturalización térmica para la proteína salvaje y las variantes 19 y 17, en presencia de 0 M NaCl (ΔG_{total} , panel superior), 2 M NaCl ($\Delta G_{no\ polar}$, panel central), y la diferencia entre ambas ($\Delta G_{coulomb}$, panel inferior). (c) Espectros de 1D-RMN registrados para el homeodominio de HESX-1 salvaje y las variantes 19 y 17. En las tres variantes se identifica fácilmente la señal de L16, de desplazamiento químico inusual y característico de homeodominios.

qué la combinación presente en la variante **1** (V19/L26/I30) no se da en ningún homeodominio natural. Parece claro, por tanto, que una cierta especificidad del empaquetamiento entre los residuos 19, 26 y 30 es crítica para la estabilidad del plegamiento del homeodominio.

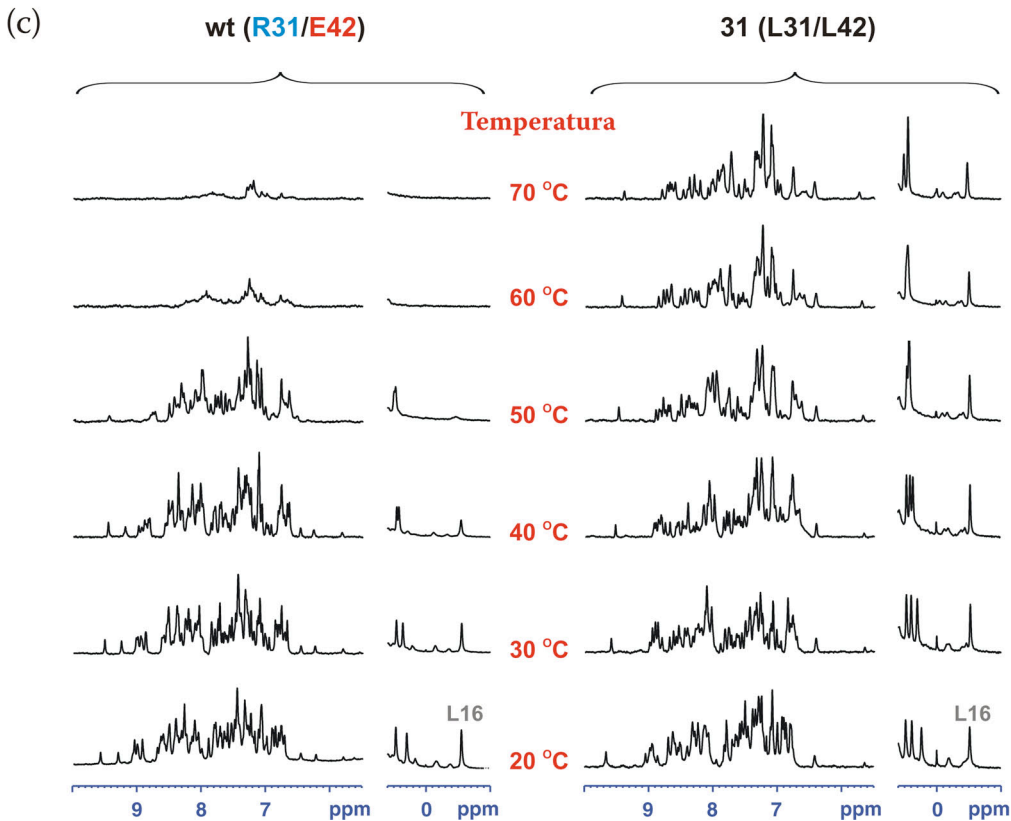
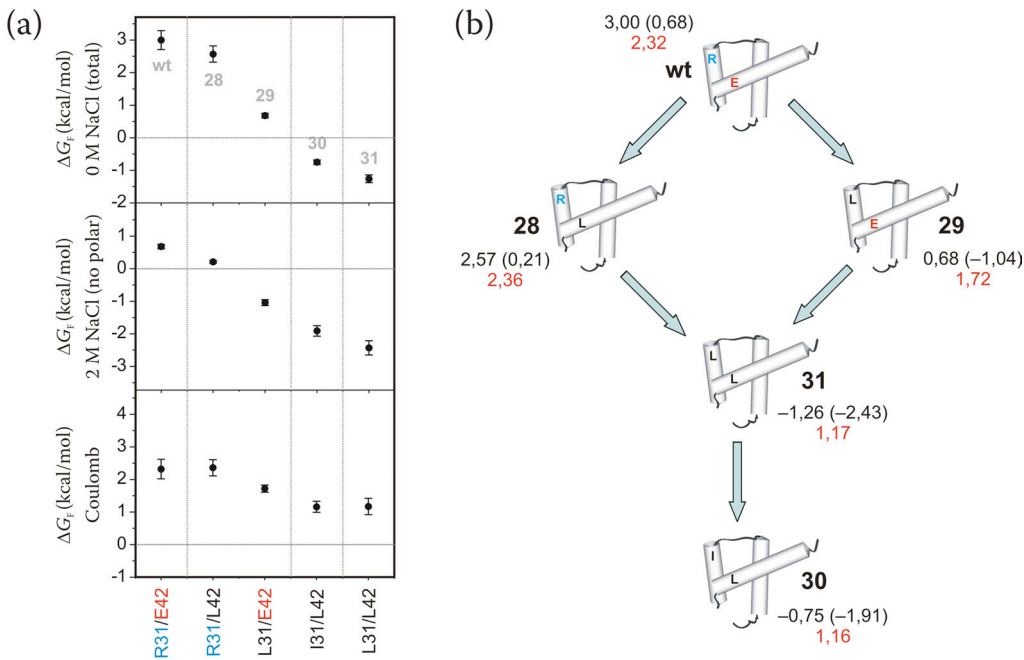
3.1.6. Puentes salinos frente a pares hidrófobos en las posiciones 31-42

Con objeto de analizar el papel del par 31-42 en la estructura y la estabilidad del homeodominio, preparamos otro grupo de mutantes simples y dobles, que fueron sujetos a experimentos de desnaturalización tanto térmica como química. Los resultados obtenidos se recogen en la [Tabla 3.1](#) y se ilustran en las [Figuras 3.10](#) y [3.11](#).

Se puede observar que la sustitución individual E42L (variante **28** [R31/L42]) incrementa ligeramente la estabilidad de la proteína. Una estabilización mucho más grande (de 2,32 kcal/mol a 77 °C) se consigue con la mutación simple R31L (variante **29** [L31/E42]). Ambos resultados son inesperados dada la localización parcialmente expuesta de los residuos 42 y 31, y teniendo en cuenta que cada una de estas mutaciones rompe el puente salino R31-E42 presente en la proteína salvaje, y observado además en la mayor parte de los homeodominios.

Combinando ambas mutaciones mediante la doble sustitución R31L/E42L, obtuvimos **31** (L31/L42), que resultó ser una versión termófila del homeodominio de HESX-1. Así, dicha variante destacó por ser 4 kcal/mol más estable que el polipéptido salvaje, presentando temperaturas de desnaturalización en el rango de 88-96 °C, dependiendo de las condiciones experimentales (ver [Tabla 3.1](#) y [Figuras 3.10](#) y [3.11](#)). Los resultados obtenidos demuestran que

FIGURA 3.10 (PÁGINA SIGUIENTE): (a) Representación gráfica de los valores de ΔG_F a 77 °C medidos a partir de experimentos de desnaturalización térmica para las variantes del homeodominio de HESX-1 con sustituciones simples o dobles en las posiciones 31-42, en presencia de 0 M NaCl (ΔG_{total} , panel superior) y 2 M NaCl ($\Delta G_{no polar}$, panel central) junto con la diferencia entre ambos valores ($\Delta G_{coulomb}$, panel inferior). (b) Representación esquemática de las variantes analizadas. Los residuos cargados negativamente se muestran en rojo, y los cargados positivamente en azul. Se dan los valores numéricos de ΔG_{total} , $\Delta G_{no polar}$ (entre paréntesis) y $\Delta G_{coulomb}$ (en rojo) a 77 °C para cada variante, obtenidos a partir de experimentos de desnaturalización térmica. Los detalles experimentales se recogen en la [Tabla 3.1](#) y en la [Figura 3.6](#). Las energías libres se dan en kcal/mol. (c) Experimentos de desnaturalización térmica monitorizados por 1D-RMN. Los espectros obtenidos son consistentes con los datos de dicroísmo circular y demuestran que, en contraste con el homeodominio salvaje (a la izquierda), la variante **31** (a la derecha) permanece plegada por encima de 70 °C. Los espectros no indican agregación proteica o heterogeneidad estructural para el doble mutante **31**.



un puente salino conectando los residuos 31 y 42 es altamente desestabilizante con respecto al par alifático L31-L42. El origen de este efecto se puede entender examinando los datos termodinámicos en detalle. De acuerdo con éstos, pese a la rotura de la interacción R31-E42, la sustitución de ambas cadenas laterales por residuos alifáticos resulta en una mejora del balance electrostático del homeodominio, como se puede deducir de los términos $\Delta G_{\text{coulomb}}$. Además, estas mutaciones conducen a unas contribuciones $\Delta G_{\text{no polar}}$ extraordinariamente favorables respecto a la proteína salvaje, lo que sugiere que los homeodominios mutantes presentan un empaquetamiento de las cadenas laterales mucho más optimizado. De hecho, las diferencias de ΔG_F a 77 °C entre las cuatro variantes a 0 M NaCl se mantienen en gran medida a 2 M NaCl, como se puede comprobar en la representación de la [Figura 3.10a](#).

En principio, $\Delta G_{\text{no polar}}$ debería ser bastante sensible al empaquetamiento de la cadena lateral alifática en posición 31. Para comprobar si esto es así, el residuo de leucina presente en la variante **31** (L31/L42) fue reemplazado por una isoleucina para dar el mutante **30** (I31/L42). Como cabría esperar (ver [Tabla 3.1](#) y [Figura 3.10, a y b](#)), los homeodominios **30** y **31** presentan idénticas contribuciones $\Delta G_{\text{coulomb}}$ a 77 °C. En cambio, la contribución independiente de la sal ($\Delta G_{\text{no polar}}$) es menos favorable en la variante **30**. La desestabilización observada podría reflejar parcialmente la menor propensión intrínseca de hélice de la isoleucina con respecto a la leucina (en torno a 0,40 kcal/mol [O'Neil y DeGrado, 1990; Chakraborty *et al.*, 1994; Pace y Scholtz, 1998]). Sin embargo, también sugiere que, en este caso concreto, la cadena lateral de la leucina establece interacciones de empaquetamiento más optimizadas dentro del núcleo hidrófobo de la proteína que las que puede establecer la isoleucina. Si bien se trata de un único dato, este resultado contrasta con estudios previos, según los cuales en proteínas termoestables se aprecia una mayor frecuencia de aparición de isoleucinas con respecto a leucinas (Kumar y Nussinov, 2001; Pack y Yoo, 2004). En cualquier caso, tanto **30** como **31** presentan valores de ΔG_F , así como de T_m , significativamente superiores a los deducidos para la proteína salvaje, lo que permite concluir que puentes alifáticos conectando las posiciones 31 y 42 (presentes en una mínima fracción de homeodominios) son claramente estabilizantes con respecto al puente salino R31-E42 (altamente conservado en homeodominios).

Estudios previos han mostrado que un incremento en el carácter hidrófobo de una proteína puede, en algunos casos, favorecer la agregación de la misma en disolución (Dill *et al.*, 1995; Marshall y Mayo, 2001). Teniendo en cuenta las localizaciones parcialmente expuestas de los residuos 31 y 42, se podría considerar la hipótesis de que uno de los papeles para la altamente conservada interacción R31-E42 es proteger la proteína contra este efecto. Con objeto de analizar esta posibilidad, registramos espectros de 1D-RMN para el

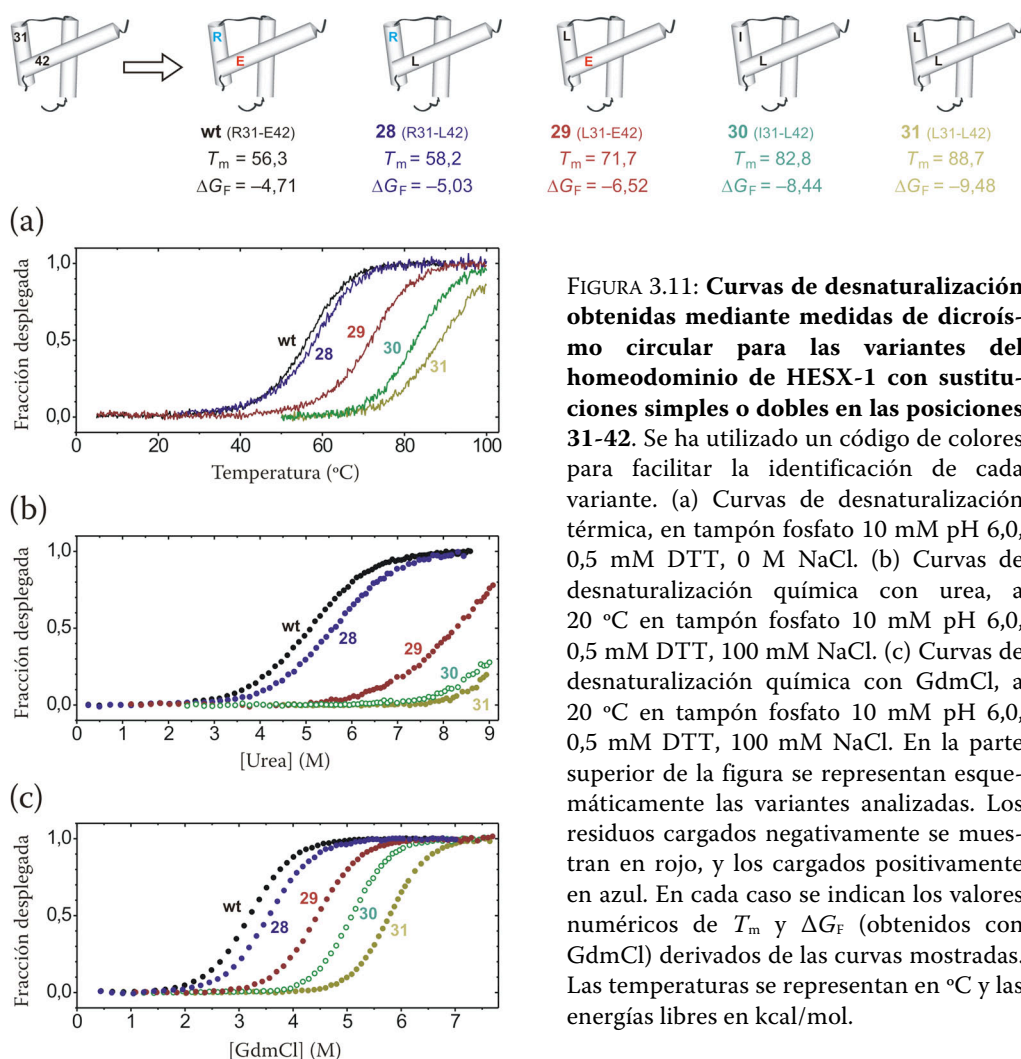


FIGURA 3.11: Curvas de desnaturalización obtenidas mediante medidas de dicróismo circular para las variantes del homeodominio de HESX-1 con sustituciones simples o dobles en las posiciones 31-42. Se ha utilizado un código de colores para facilitar la identificación de cada variante. (a) Curvas de desnaturalización térmica, en tampón fosfato 10 mM pH 6,0, 0,5 mM DTT, 0 M NaCl. (b) Curvas de desnaturalización química con urea, a 20 °C en tampón fosfato 10 mM pH 6,0, 0,5 mM DTT, 100 mM NaCl. (c) Curvas de desnaturalización química con GdmCl, a 20 °C en tampón fosfato 10 mM pH 6,0, 0,5 mM DTT, 100 mM NaCl. En la parte superior de la figura se representan esquemáticamente las variantes analizadas. Los residuos cargados negativamente se muestran en rojo, y los cargados positivamente en azul. En cada caso se indican los valores numéricos de T_m y ΔG_F (obtenidos con GdmCl) derivados de las curvas mostradas. Las temperaturas se representan en °C y las energías libres en kcal/mol.

homeodominio salvaje y para la variante **31** (L31/L42) a diferentes temperaturas (Figura 3.10c). Puede observarse como la proteína termoestable exhibe una buena dispersión en los espectros 1D de protón, rasgo característico de las proteínas bien plegadas. Además, la ausencia de un ensanchamiento significativo de las señales descarta la existencia de una agregación significativa bajo las condiciones experimentales empleadas (con una concentración de la proteína de 500 μ M). Finalmente, la variación de los datos espectroscópicos con la temperatura está en consonancia con los experimentos de desnaturalización de dicróismo circular, confirmando un incremento notable en la estabilidad de **31**. Así, puede observarse que las señales de HESX-1 salvaje se

debilitan mucho por encima de 50 °C como consecuencia de su desplegamiento, mientras que el doble mutante **31** se mantiene correctamente plegado, dando señales finas y claras incluso a 70 °C.

En conclusión, cadenas laterales alifáticas en posiciones 31 y 42 inducen una notable estabilización del homeodominio sin promover una agregación significativa incluso a altas concentraciones de proteína.

3.1.7. Determinación de la estructura 3D en disolución de la variante termoestable 31 (mutante R31L/E42L) mediante RMN

Con objeto de determinar la influencia de las sustituciones R31L/E42L en la conformación del homeodominio y analizar las bases moleculares de su termoestabilidad, decidimos abordar el estudio estructural de la variante **31** (L31/L42) empleando métodos de resonancia magnética nuclear (RMN).

El primer paso en la elucidación estructural de proteínas por RMN es la asignación completa de su espectro protónico. En el caso de la variante **31**, dicha asignación se basó en la del polipéptido salvaje previamente analizado en nuestro grupo de trabajo, lo cual agilizó el proceso en buena medida. De hecho, la simple comparación de los espectros obtenidos para **31** con los correspondientes a la proteína salvaje permitió identificar de un modo sencillo aquellos

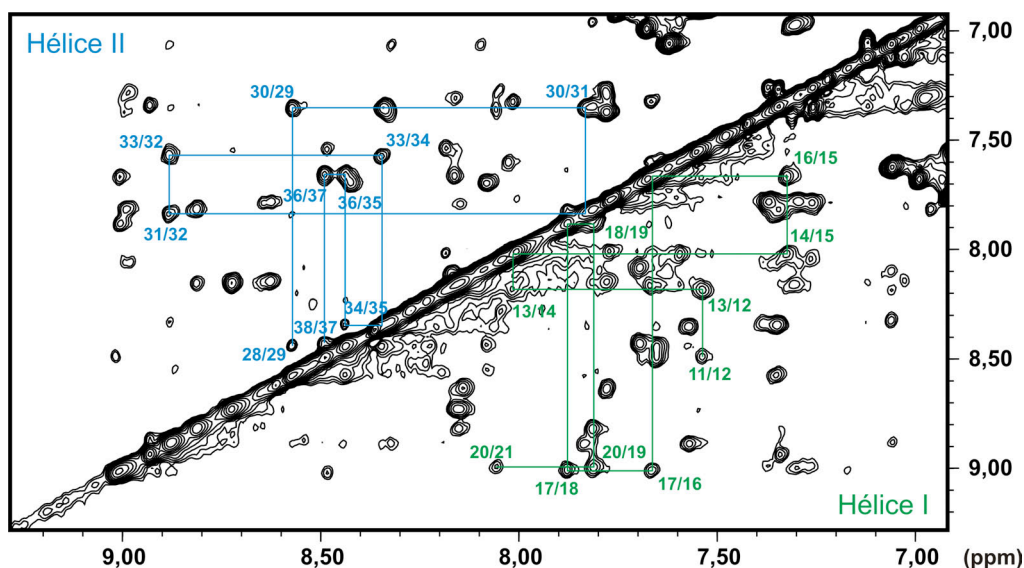


FIGURA 3.12: Región HN-HN del espectro 2D NOESY medido en H₂O para la variante termoestable **31**. En ella se indican las conexiones secuenciales entre protones amídicos pertenecientes a las hélices I y II, como ejemplo del proceso de asignación.

sistemas de espín correspondientes a zonas de la proteína no afectadas por la mutación. La asignación completa se consiguió siguiendo el método estándar desarrollado por Wüthrich y colaboradores (Wüthrich, 1986) basado en la conexión secuencial de los diferentes sistemas de espín a partir de picos NOE en experimentos 2D NOESY medidos en H₂O (Figura 3.12). Dicha información se complementó con la derivada de experimentos homonucleares estándar DQF-COSY y 2D TOCSY adquiridos tanto en H₂O como en D₂O. De este modo, se obtuvo la práctica totalidad de las asignaciones con la excepción de las correspondientes a los residuos 1 y 2, así como las del metilo ϵ de M67. Las dificultades para asignar dichas posiciones tienen su origen en su localización en zonas extremadamente flexibles del polipéptido (como corresponde a los extremos N- y C-terminal). En la Tabla 3.2 se dan los desplazamientos químicos de protón obtenidos.

Un primer diagnóstico de las alteraciones estructurales presentes en el mutante **31** pudo obtenerse comparando sus asignaciones con las correspondientes a la proteína salvaje. Este análisis descarta la existencia de mayores reorganizaciones estructurales como consecuencia de las mutaciones. En la Figura 3.13 se muestran las perturbaciones de desplazamiento químico inducidas por las sustituciones R31L/E42L. Puede apreciarse como, aparte de las propias posiciones mutadas, existen tres zonas especialmente afectadas: el extremo C-terminal de la hélice I, la primera mitad de la hélice III, y el lazo I-II (el origen de estas alteraciones se discutirá más adelante). Además, se observa, en general, una mayor variación en desplazamiento químico en aquellas

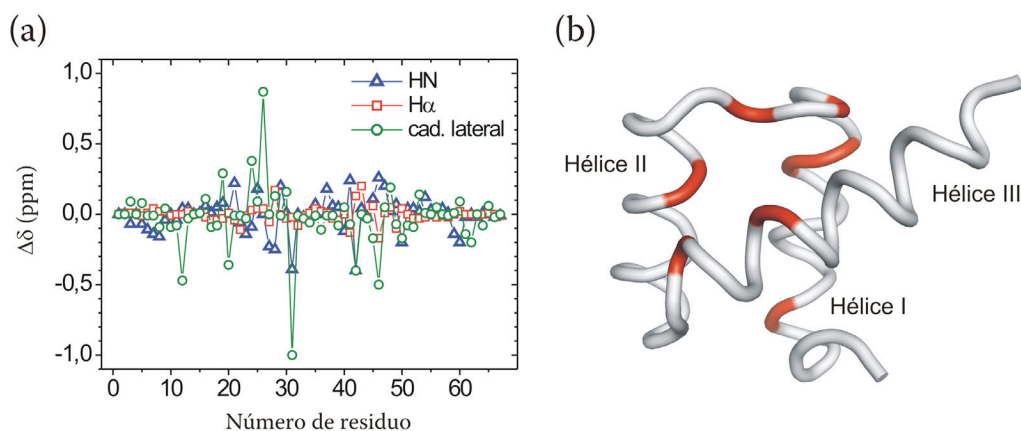


FIGURA 3.13: (a) Diferencias de desplazamiento químico entre el homeodominio de HESX-1 salvaje y la variante **31**, observadas para los protones HN (azul), H α (rojo) y de cadena lateral (verde). (b) Representación simplificada de la estructura tridimensional de la proteína salvaje. Aquellos residuos con $\Delta\delta > 0,25$ ppm se muestran en rojo.

TABLA 3.2: Desplazamientos químicos de protón (ppm) de la variante termoestable del homeo-
dominio de HESX-1 (31) obtenidos a 25 °C en tampón fosfato 10 mM pH 6,0. Se utilizó TSP
como referencia interna

Residuo	HN	Hα	Hβ	Otros
Gly 1	-	-	-	-
Arg 2	-	4,39	-	-
Arg 3	8,58	4,63	1,94, 1,86	Hy 1,77, 1,70 Hδ 3,22, 3,21
Pro 4	-	4,45	2,32, 1,88	Hy 2,04, 2,02 Hδ 3,85, 3,66
Arg 5	8,56	4,36	1,93, 1,82	Hy 1,71, 1,66 Hδ 3,22, 3,21
Thr 6	8,19	4,26	4,08	Hy 1,19
Ala 7	8,23	4,34	1,36	-
Phe 8	8,32	4,96	3,28, 2,90	Hδ 7,06 He 7,12 Hζ 7,03
Thr 9	8,88	4,50	4,79	Hy 1,33
Gln 10	9,02	4,00	2,19, 2,08	Hy 2,48, 2,47 He 7,76, 6,92
Asn 11	8,48	4,41	2,78, 2,68	Hδ 7,68, 6,99
Gln 12	7,54	3,76	2,10, 1,41	Hy 2,27, 1,79 He 7,49, 6,78
Ile 13	8,19	3,13	1,87	Hy1 1,69, 1,10 Hy2 1,00 Hδ 0,83
Glu 14	8,01	3,92	2,12, 2,03	Hy 2,43, 2,24
Val 15	7,33	3,64	2,08	Hy 1,04, 0,85
Leu 16	7,67	3,57	0,56, -0,46	Hy 1,23 Hδ 0,47, -0,54
Glu 17	9,00	4,26	2,17, 2,01	Hy 2,46, 2,44
Asn 18	7,88	4,55	3,06, 2,85	Hδ 7,62, 7,06
Val 19	7,81	3,88	2,49	Hy 1,20, 1,44
Phe 20	8,99	4,32	3,44, 3,30	Hδ 7,30 He 6,98 Hζ 6,97
Arg 21	8,06	3,96	2,03, 1,95	Hy 1,93, 1,79 Hδ 3,31, 3,30
Val 22	7,35	4,17	2,29	Hy 1,06, 0,97
Asn 23	8,34	4,60	2,78, 2,63	Hδ 7,48, 6,76
Cys 24	8,17	4,13	2,91, 2,64	-
Tyr 25	8,10	4,66	3,21, 2,66	Hδ 7,06 He 6,82
Pro 26	-	4,21	1,78, 1,53	Hy 1,17, 0,97 Hδ 3,31, 2,96
Gly 27	8,23	4,15, 3,93	-	-
Ile 28	8,44	3,94	1,94	Hy1 1,46, 1,45 Hy2 0,97 Hδ 0,96
Asp 29	8,57	4,34	2,72, 2,58	-
Ile 30	7,36	3,88	1,97	Hy1 1,57, 1,28 Hy2 0,87 Hδ 0,86
Leu 31	7,84	3,90	1,96, 1,25	Hy 1,55 Hδ 0,74, 0,39
Glu 32	8,88	4,02	2,17, 2,01	Hy 2,54, 2,26
Asp 33	7,57	4,49	2,86, 2,80	-
Leu 34	8,35	3,92	2,16, 1,22	Hy 1,97 Hδ 0,96, 0,83
Ala 35	8,44	3,90	1,58	-
Gln 36	7,66	4,17	2,27, 2,25	Hy 2,60, 2,47 He 7,55, 6,90
Lys 37	8,49	4,03	1,93, 1,81	Hy 1,59, 1,43 Hδ 1,61, 1,60 He 2,98, 2,95
Leu 38	8,42	4,36	1,69, 1,40	Hy 1,81 Hδ 0,81, 0,67
Asn 39	7,69	4,37	3,18, 2,71	Hδ 7,48, 6,82
Leu 40	8,08	4,72	1,29, 1,27	Hy 1,42 Hδ 0,84, 0,30
Glu 41	8,54	4,21	1,93, 1,92	Hy 2,48, 2,31
Leu 42	8,87	3,88	1,85, 1,64	Hy 1,58 Hδ 0,94, 0,88
Asp 43	8,93	4,54	2,74, 2,71	-
Arg 44	7,34	4,22	2,23, 2,17	Hy 2,01, 1,95 Hδ 3,32, 3,30
Ile 45	7,81	3,96	2,28	Hy1 1,70, 1,02 Hy2 1,21 Hδ 0,78
Gln 46	8,81	4,24	2,38, 2,19	Hy 2,52, 2,33 He 7,23, 6,87
Ile 47	8,15	3,85	2,02	Hy1 1,79, 1,30 Hy2 1,00 Hδ 0,88
Trp 48	8,14	4,04	3,68, 3,31	Hδ1 6,88 He1 9,60 He3 6,40 Hζ2 7,17 Hζ3 5,64 Hη2 6,40
Phe 49	8,63	3,80	3,38, 3,31	Hδ 7,78 He 7,37 Hζ 7,26
Gln 50	8,15	3,96	2,30, 2,20	Hy 2,59, 2,49 He 7,54, 6,90
Asn 51	8,72	4,34	2,71, 2,51	Hδ 7,55, 6,90

CONTINÚA EN LA PÁGINA SIGUIENTE

TABLA 3.2 (CONT.)

Residuo	HN	H α	H β	Otros
Arg 52	8,16	3,49	0,74, -0,13	Hy -0,15, -0,39 H δ 2,33, 1,95
Arg 53	8,16	4,10	2,18, 1,85	Hy 2,32, 1,51 H δ 2,74, 2,51
Ala 54	7,67	4,17	1,51	-
Lys 55	7,60	4,00	1,80, 1,74	Hy 1,45, 1,34 H δ 1,52, 1,51 He 2,86, 2,85
Leu 56	7,60	4,27	1,89, 1,88	Hy 1,77 H δ 1,07, 1,02
Lys 57	8,03	4,17	1,91, 1,87	Hy 1,58, 1,47 H δ 1,74, 1,72 He 3,02, 3,01
Arg 58	8,00	4,26	1,90, 1,87	Hy 1,73, 1,67 H δ 3,21, 3,21
Ser 59	8,01	4,40	3,95, 3,91	-
His 60	8,13	4,72	3,32, 3,30	H δ 7,26 He 8,35
Arg 61	8,27	4,29	1,83, 1,81	Hy 1,65, 1,63
Glu 62	8,52	4,26	2,01, 1,98	Hy 2,12, 2,11
Ser 63	8,24	4,37	3,91, 3,84	-
Gln 64	8,28	4,24	1,96, 1,89	Hy 2,15, 2,13 He 7,43, 6,87
Phe 65	8,05	4,66	3,21, 2,99	H δ 7,26 He 7,37 H ζ 7,35
Leu 66	8,01	4,37	1,63, 1,59	Hy 1,56 H δ 0,91, 0,86
Met 67	7,77	4,24	2,09, 1,95	Hy 2,54, 2,48

posiciones orientadas hacia el núcleo hidrófobo, lo que pone de manifiesto la reorganización del mismo motivada por las mutaciones. Nuestro análisis revela también cambios de desplazamiento químico pronunciados en los residuos Q12, V19 y Q46, cuyo origen se discutirá al final del presente apartado. Finalmente, es interesante destacar que se mantienen los desplazamientos químicos inusuales típicos del plegamiento de homeodominio, como son el metilo $\delta 2$ de L16, que aparece desplazado a campo alto hasta $-0,54$ ppm, o los protones aromáticos He3, H ζ 3 y H η 2 de W48, desplazados a campo alto por debajo de 6,50 ppm.

A pesar de los pequeños ajustes conformacionales sugeridos por el análisis recién descrito, una inspección detallada de los picos de cruce NOE a media y larga distancia demuestra que la estructura secundaria y terciaria del homeodominio permanece fundamentalmente inalterada. Así, un patrón de NOEs secuenciales y de media distancia característicos de hélice α (NOEs secuenciales HN-HN [i , $i+1$] intensos o medios, y presencia de NOEs H α -HN [i , $i+3$], H α -HN [i , $i+4$], H α -H β [i , $i+3$] y H α -H β [i , $i+4$]) demuestra la persistencia de los tres segmentos de hélice α presentes en la proteína salvaje y típicos de homeodominios, que cubren los tramos 10 a 22, 28 a 38 y 42 a 56. Los residuos 56 a 60 de la región C-terminal, pese a estar algo desordenados, no presentan una estructura indefinida, a juzgar por la presencia de NOEs débiles que conectan las posiciones i , $i+3$ e i , $i+4$ en esta región, que sugiere la existencia de una estructura helicoidal poco poblada.

Los tres segmentos de hélice α empaquetan de un modo análogo a la proteína salvaje, a juzgar por varios NOEs inequívocos de larga distancia que

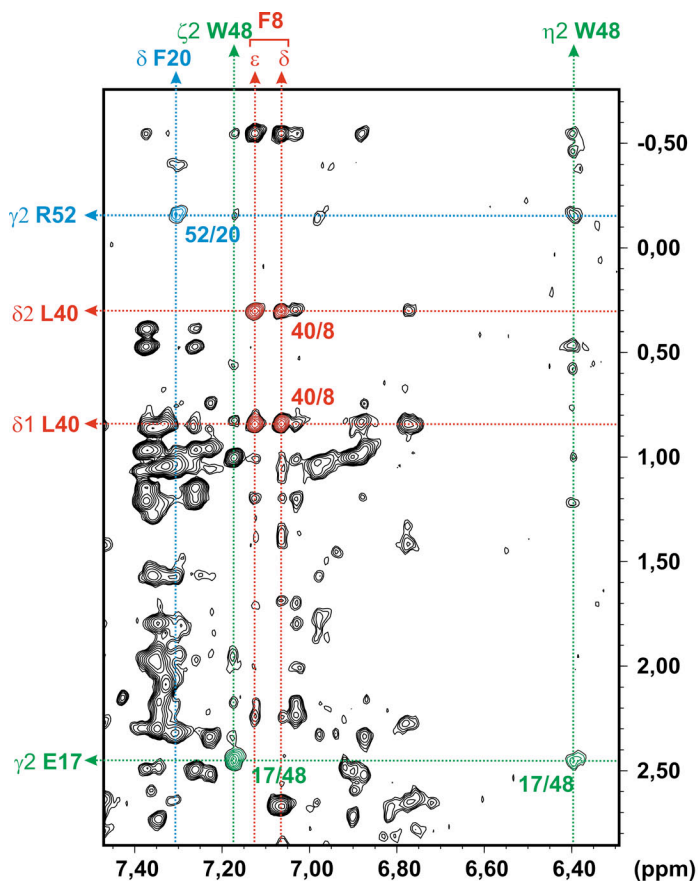


FIGURA 3.14: Región HN-alifático correspondiente a un espectro 2D NOESY medido en H₂O para la variante 31. En ella se destacan diversos picos de cruce entre residuos alejados en la secuencia y, por tanto, característicos de la estructura terciaria del polipéptido. A la derecha se representa esquemáticamente la posición de esos residuos en la estructura del homeodominio.

conectan regiones alejadas en la secuencia, tales como los que se dan entre L40 y F8, entre R52 y F20, o entre E17 y W48 (ver [Figura 3.14](#)).

A partir de la inspección detallada de diversos experimentos NOESY adquiridos tanto en agua ligera como deuterada, se asignaron 703 NOEs no ambiguos de los cuales 358 corresponden a contactos a media y larga distancia (ver [Tabla 3.3](#)). Para las regiones N-terminal (residuos 1 a 7) y C-terminal (residuos 56 a 67) se midieron un escaso número de NOEs, como resultado de su elevada movilidad (rasgo compartido con el polipéptido salvaje).

Los NOEs obtenidos de esta forma se tradujeron de un modo cualitativo a restricciones de distancia (ver la parte de procedimientos experimentales) y se introdujeron en cálculos de algoritmo de distancias con el programa DYANA (Guntert *et al.*, 1997). Las 30 mejores estructuras obtenidas (en términos de función diana y violaciones) se sometieron a un refinamiento posterior mediante cálculos de dinámica molecular restringida, empleando el campo de fuerzas AMBER 5.0. Éstos se realizaron en presencia de disolvente explícito con

TABLA 3.3: Estadísticas correspondientes al cálculo de la estructura del mutante termófilo **31** por RMN, en diferentes etapas de refinamiento. Los valores de RMSD y de violaciones se dan en Å

Restricciones de distancia de RMN	
Restricciones de distancia de NOE totales	703
Intrarresiduo	164
Secuenciales	181
Media y larga distancia	358
Puentes de hidrógeno	11
Estadísticas de la estructura	
Función diana de DYANA (30 estructuras)	0,21 ± 0,02
RMSD del esqueleto por DYANA (residuos 8 a 55, 30 estructuras)	0,47 ± 0,09
RMSD de átomos pesados por DYANA (residuos 8 a 55, 30 estructuras)	1,27 ± 0,12
RMSD del esqueleto por AMBER (residuos 8 a 55, 25 estructuras)	0,48 ± 0,07
RMSD de átomos pesados por AMBER (residuos 8 a 55, 25 estructuras)	1,36 ± 0,14
Violaciones de restricciones de distancia (>0,2 Å)	
Suma promedio	3,1
Máxima promedio	0,25
Desviaciones respecto a la geometría ideal	
Distancias de enlace	0,023 ± 3,3·10 ⁻⁴
Ángulos de enlace	2,24 ± 0,04
Estadísticas de Ramachandran	
Regiones más favorables	86,3%
Adicionalmente permitidas	13,7%
Generosamente permitidas	0,0%

objeto de impedir un empaquetamiento artefactual de los extremos N- y C-terminal (fundamentalmente desordenados) sobre el núcleo estructurado de la proteína. De este modo, se obtuvieron 25 estructuras caracterizadas por bajos valores de energía y virtualmente libres de violaciones. En la [Tabla 3.3](#) se muestran las estadísticas correspondientes a las estructuras generadas en cada una de las etapas de refinamiento.

La [Figura 3.15](#) muestra el conjunto de las 25 estructuras finales superpuestas, correspondientes al mutante **31** en disolución, de acuerdo a los datos de RMN. Puede apreciarse como, con excepción de los extremos N- y C-terminal, la definición es óptima, tanto en lo que respecta a la cadena principal (RMSD 0,48 ± 0,07 Å) como a los átomos pesados (RMSD 1,36 ± 0,14 Å, ver [Tabla 3.3](#)). En general, presentan desviaciones geométricas muy pequeñas con respecto a los valores ideales, así como contactos no covalentes razonables. De hecho, un análisis de la distribución de Ramachandran demuestra que los ángulos de torsión se encuentran en su mayoría en las regiones más favorables ([Tabla 3.3](#)).

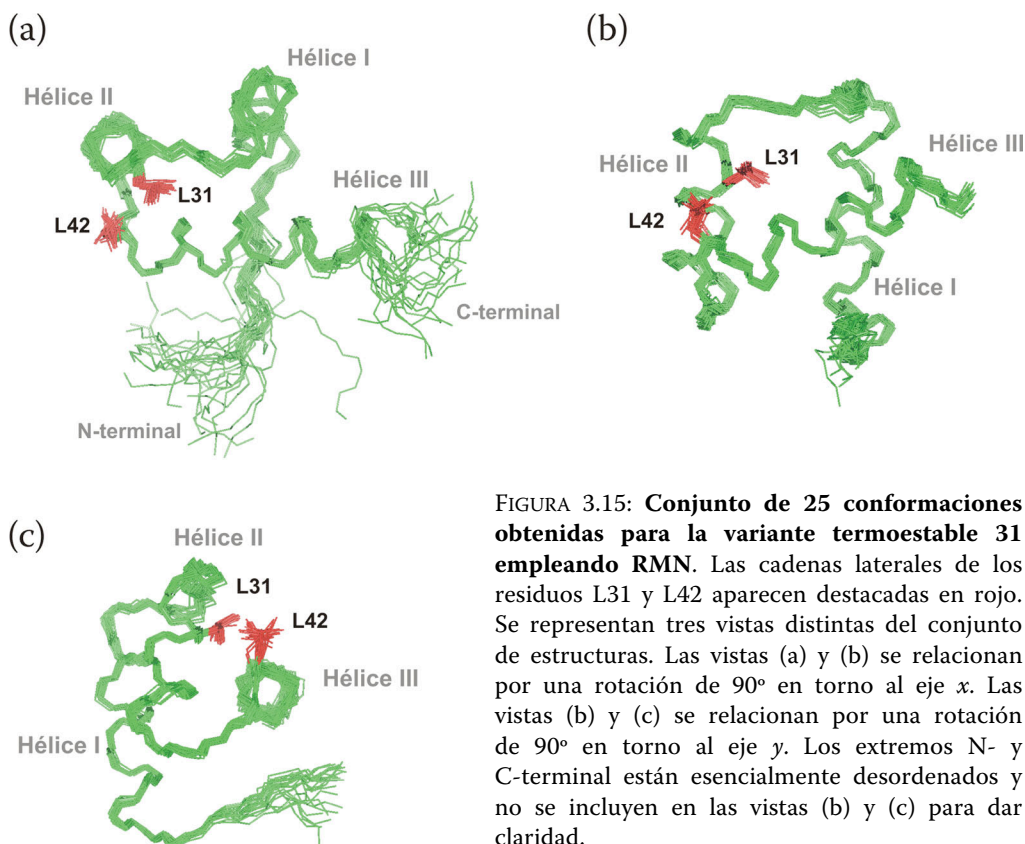


FIGURA 3.15: **Conjunto de 25 conformaciones obtenidas para la variante termoestable 31 empleando RMN.** Las cadenas laterales de los residuos L31 y L42 aparecen destacadas en rojo. Se representan tres vistas distintas del conjunto de estructuras. Las vistas (a) y (b) se relacionan por una rotación de 90° en torno al eje x . Las vistas (b) y (c) se relacionan por una rotación de 90° en torno al eje y . Los extremos N- y C-terminal están esencialmente desordenados y no se incluyen en las vistas (b) y (c) para dar claridad.

Puede apreciarse como la definición de las cadenas laterales que integran el núcleo hidrófobo es, en general, excelente. Dichas cadenas establecen un patrón de interacciones muy característico de homeodominios, como hemos descrito en la introducción (Figura 3.16). Es de destacar que los contactos observados son plenamente consistentes con los desplazamientos químicos más inusuales incluidos en nuestra asignación de la variante 31 que señalábamos anteriormente, y también presentes en la proteína salvaje (de la Mata *et al.*, 2002) así como en otros homeodominios (Güntert *et al.*, 1991; Tsao *et al.*, 1995; Mayor *et al.*, 2003). Así por ejemplo, el inusual desplazamiento a campo alto del metilo $\delta 2$ de L16 es consecuencia de su apilamiento sobre el anillo aromático de W48. Del mismo modo, el desplazamiento a campo alto de los protones aromáticos H $\epsilon 3$, H $\zeta 3$ y H $\eta 2$ de W48 se explica por su interacción (tipo T) con F49. Estos tres residuos (L16, W48 y F49) aparecen conservados en la práctica totalidad de las secuencias de homeodominios conocidas. Los contactos que establecen se reflejan en numerosos NOEs (Figura 3.16, b y c).

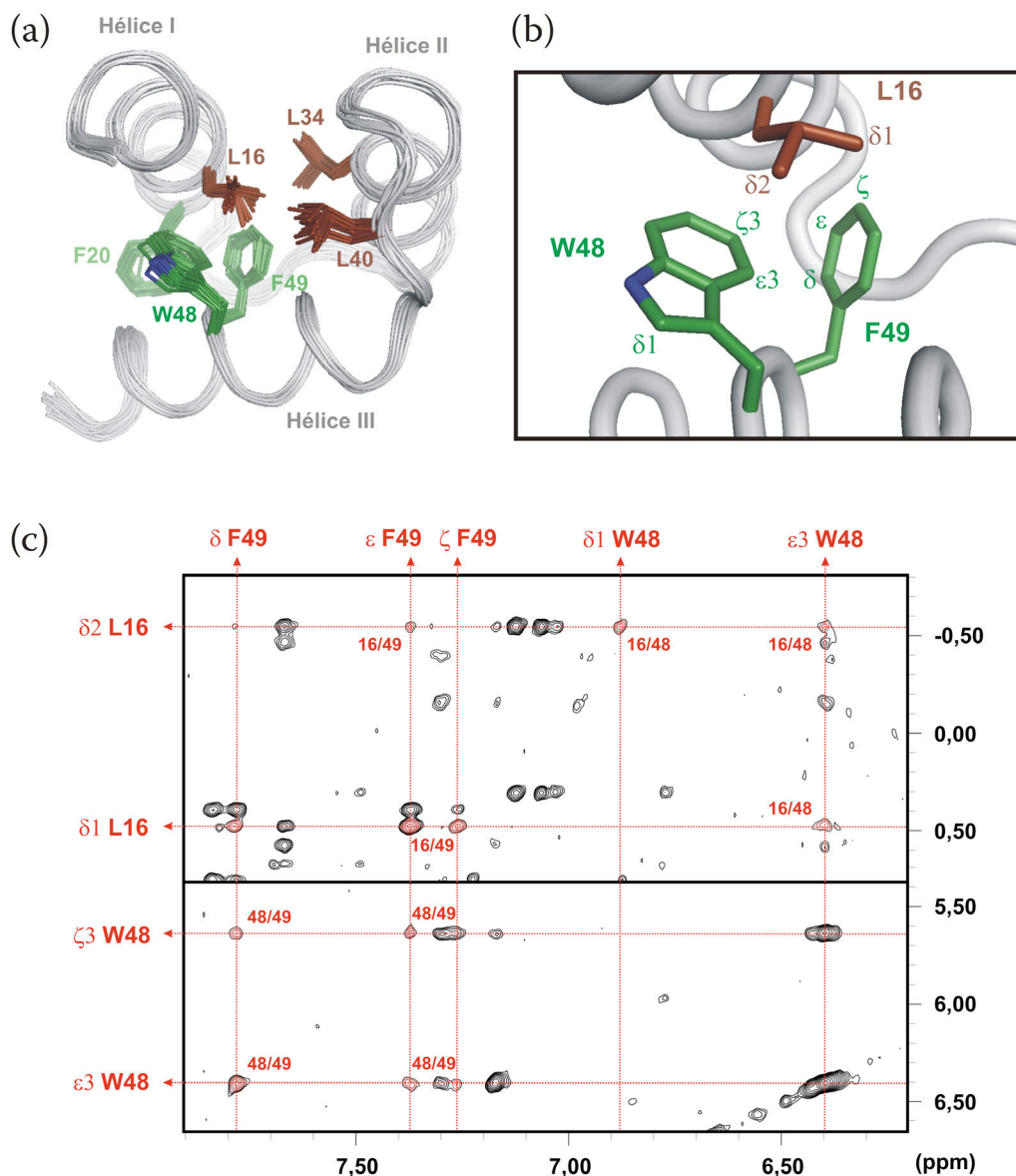
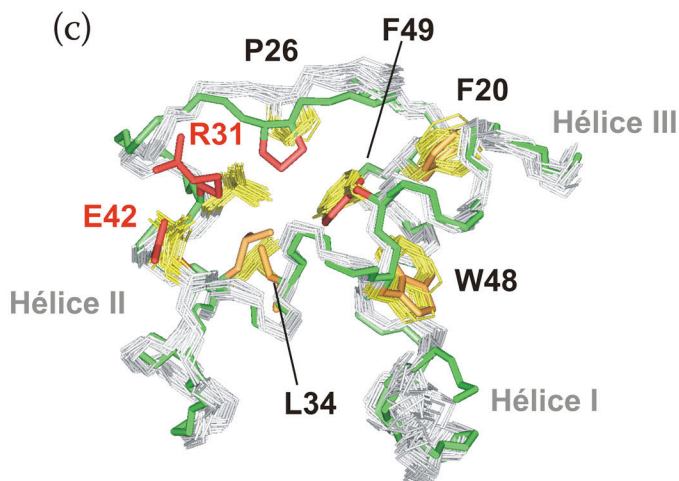
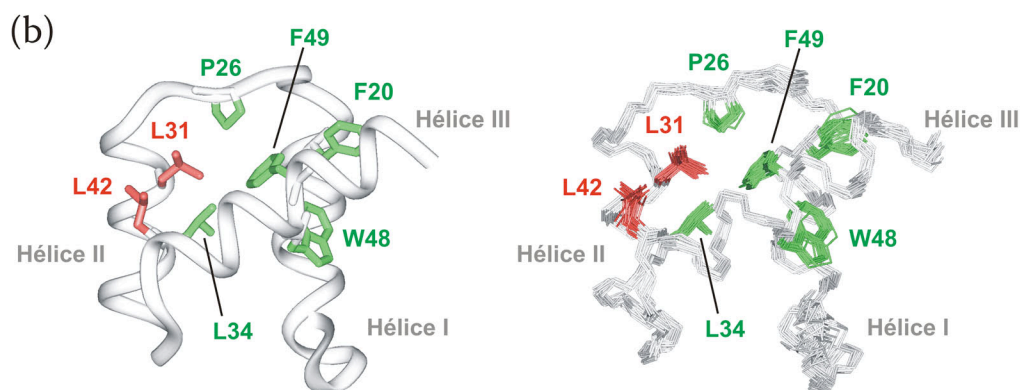
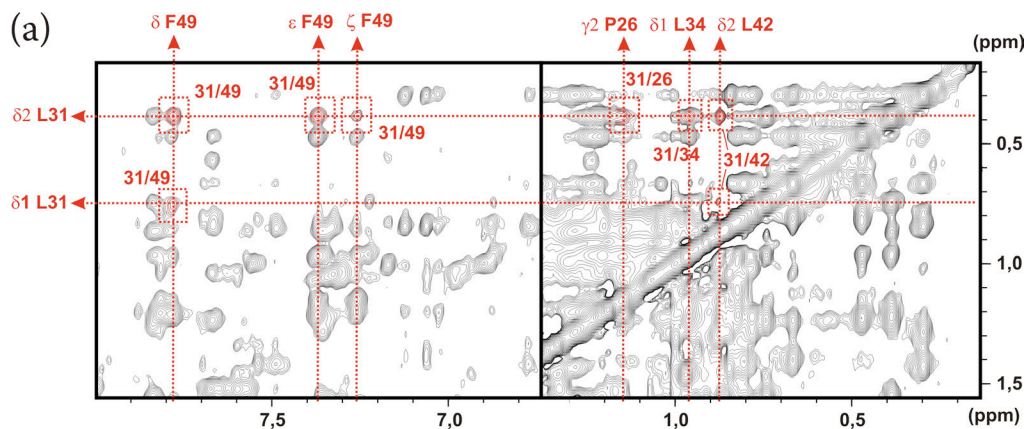


FIGURA 3.16: (a) Superposición de 25 estructuras de RMN de la variante **31**, en la que se destacan algunos de los residuos aromáticos (verde) y alifáticos (marrón) más importantes en la formación del núcleo hidrófobo de homeodominios. (b) Detalle de una única estructura representativa, que muestra el patrón de interacciones más característico y conservado del núcleo hidrófobo de homeodominios, entre los residuos L16 (marrón), W48 y F49 (ambos en verde). (c) Se muestran dos regiones diferentes de un espectro 2D NOESY medido en H₂O para la variante **31**. En ellas se destacan diversos picos de cruce que involucran los residuos anteriormente mencionados.

Una inspección superficial de la estructura de **31** no permite apreciar cambios significativos con respecto al homeodominio salvaje. De hecho, la superposición del esqueleto peptídico de ambas proteínas entre los residuos 8 y 55 arroja un valor de RMSD de 0,51 Å. Sin embargo, un análisis más detallado revela los pequeños ajustes estructurales que acompañan a la doble mutación. Dichos ajustes son de naturaleza local e involucran fundamentalmente la posición 31, así como diversos residuos del lazo I-II. La [Figura 3.17c](#) muestra la superposición entre las estructuras derivadas por RMN para la proteína salvaje y el mutante termoestable. Puede observarse como, en el primer caso, R31 está parcialmente expuesta al disolvente, e involucrada en un puente salino con E42. Por su parte, el residuo de prolina P26, localizado en el lazo I-II, presenta una clara interacción de apilamiento con el anillo aromático de F49 (conservada en un buen número de homeodominios). Este contacto se refleja en los desplazamientos químicos inusuales de P26 (protones Hy en 1,04/0,12 ppm). De acuerdo con los datos obtenidos, la doble sustitución R31L/E42L promueve un pequeño cambio en la orientación de la cadena lateral de L31, que se inserta en el núcleo hidrófobo de la proteína, desplazando ligeramente a P26. Como consecuencia, el apilamiento de ésta con F49 se ve claramente afectado. Esta conclusión está avalada por un buen número de NOEs que conectan las posiciones 31-26 y 31-49 ([Figura 3.17a](#)) así como por los desplazamientos químicos de P26, ahora más próximos al rango esperado para un residuo de prolina (protones Hy en 1,17/0,97 ppm, ver [Tabla 3.2](#)). En resumen, la inserción de L31 en el núcleo de la proteína tiene como consecuencia el ligero desplazamiento del lazo I-II, así como un pequeño reajuste, de naturaleza local, del empaquetamiento del núcleo hidrófobo.

A diferencia del residuo 31, la leucina en posición 42 está parcialmente expuesta al disolvente. Sin embargo, este residuo contribuye significativamente a la estabilización de la proteína a través de su interacción con L31. La estrecha proximidad entre las cadenas laterales de estos dos residuos queda de

FIGURA 3.17 (PÁGINA SIGUIENTE): (a) Se representan dos regiones diferentes de un espectro 2D NOESY homonuclear obtenido para la variante termoestable **31**. Diversos contactos clave entre la cadena lateral de L31 y los residuos F49, P26 y L42 se destacan en rojo. (b) Conjunto de 25 estructuras (a la derecha) obtenidas para la variante **31** mediante RMN. A la izquierda se muestra una única estructura representativa para facilitar su visualización. En ambos casos, el esqueleto peptídico se representa en gris, las cadenas laterales de los residuos 31 y 42 en rojo, y las de los residuos del núcleo hidrófobo 20, 26, 34, 48 y 49 en verde. (c) Conjunto de 25 estructuras de la variante **31** (el esqueleto peptídico se representa en gris y ciertas cadenas laterales clave en amarillo) superpuestas a la estructura promedio del homeodominio de HESX-1 salvaje (el esqueleto peptídico se representa en verde y ciertas cadenas laterales clave en rojo). La interacción P26-F49, observada en el homeodominio salvaje, queda parcialmente interrumpida en **31** debido al enterramiento de la cadena lateral de L31 en el núcleo hidrófobo.



manifiesto por varios NOEs (Figura 3.17a). De hecho, la estructura 3D obtenida muestra que la cadena lateral alifática de L42 reduce significativamente la superficie accesible al disolvente de L31, que queda así enterrada casi por completo en la variante **31**. Un comportamiento similar se ha observado en simulaciones de dinámica molecular para el mutante R31L del homeodominio de Paired (Iurcu-Mustata *et al.*, 2001). Curiosamente, en contraste con nuestros resultados experimentales, este estudio teórico predice un descenso en estabilidad asociado a dicha sustitución.

Otras diferencias menos llamativas entre las estructuras del homeodominio salvaje y la variante **31** serían consistentes con las diferencias en los desplazamientos químicos entre ambas proteínas que observábamos anteriormente (Figura 3.13), y podrían colaborar al aumento de estabilidad de la variante **31**. Así por ejemplo, la parte del brazo N-terminal que conecta con el inicio de la hélice I, aparece en **31** desplazada hacia fuera respecto a la proteína salvaje. Como consecuencia, el anillo aromático de F8 queda más cerca de Q12, la cual sufre de este modo un desplazamiento de todos de sus protones a campo alto, pese a ser una posición situada a más de 12 Å del punto de mutación más cercano (Figura 3.18a). Este efecto puede deberse a un cierto reajuste del núcleo aromático (formado por los residuos F49, W48 y F8, que conectan las posiciones 31 y 12), que expulsaría ligeramente a F8 de su interior. En otro ejemplo, los residuos V19 e I30 se separan levemente del núcleo hidrófobo y entre sí para dejar hueco a L31, lo que podría ser el motivo del desplazamiento a campo bajo del metilo γ_2 de V19 (Figura 3.18b). Por último, Q46 aparece en **31** más ordenada y orientada hacia el núcleo hidrófobo que en la proteína salvaje, estableciendo así contactos con L31, lo que podría justificar el desplazamiento a campo alto del NH ϵ_1 (Figura 3.18c).

En conclusión, los datos de RMN son consistentes con una reestructuración del núcleo hidrófobo en la proteína termoestable, lo que está de acuerdo con las contribuciones $\Delta G_{\text{no polar}}$ claramente estabilizantes medidas para las variantes **30** (I31/L42) y **31** (L31/L42). Este hecho, unido a un balance electrostático mejor optimizado para las proteínas mutantes, constituye la principal causa de su pronunciada estabilidad.

Por último, conviene destacar que la marcada estabilización obtenida en la variante **31** (L31/L42) es el resultado de tan solo dos mutaciones. En el contexto del diseño de proteínas, efectos comparables son, a menudo, el resultado de la introducción de un buen número de sustituciones específicamente diseñadas a tal efecto. A modo de ejemplo, utilizando el homeodominio de Engrailed como sistema de estudio, el grupo de Mayo ha demostrado en los últimos años la eficacia de los métodos computacionales para el diseño de

(a)

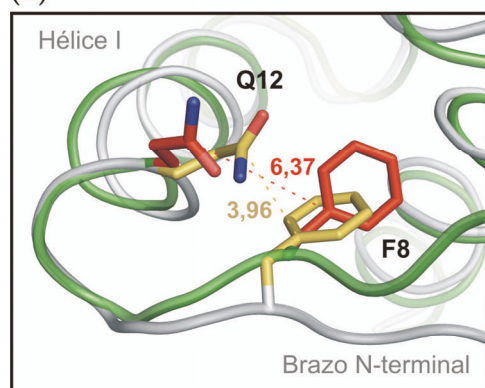
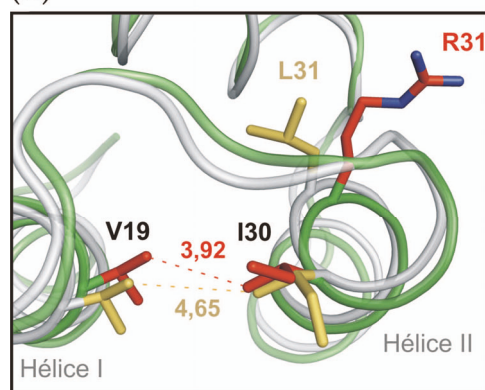
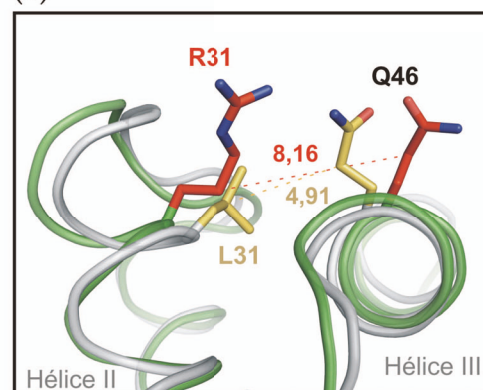


FIGURA 3.18: Ejemplos de diferencias conformacionales entre los polipéptidos salvaje y 31, en torno a residuos para los que se detectan variaciones significativas de desplazamiento químico por RMN. (a) Aproximación entre los residuos 8 y 12 en 31. (b) Separación entre los residuos 19 y 30 en 31. (c) Interacción entre los residuos 31 y 46 en 31. En los tres casos, se representa en verde y en blanco los esqueletos peptídicos, y en rojo y en amarillo las cadenas laterales correspondientes a las proteínas salvaje y variante 31, respectivamente. Se muestra, además, la distancia en Å entre pares de átomos representativos.□

(b)



(c)



proteínas termoestables. Como hemos comentado en la introducción, estos trabajos condujeron a la obtención de variantes múltiples del homeodominio con valores de T_m de 88 °C (Marshall *et al.*, 2002), >99 °C (Shah *et al.*, 2007) y 114 °C (Marshall y Mayo, 2001). Cabe mencionar que estas mutaciones conducían en algunos casos a desviaciones estructurales significativas respecto a la proteína salvaje de partida. En contraste con estos resultados, nuestro trabajo muestra que es posible la obtención de homeodominios termoestables mediante dos simples sustituciones. Estos cambios no conllevan más que mínimos ajustes estructurales (a juzgar por la estructura en disolución de la variante 31) y no promueven la agregación proteica.

3.1.8. Puentes salinos frente a pares hidrófobos en las posiciones 17-52

El puente salino 17-52, presente en el 75% de secuencias, es el más conservado en homeodominios, y su papel en la estabilidad de estos sistemas ha sido objeto de diversos estudios en el pasado. Así por ejemplo, como hemos adelantado en la introducción, la variante de Engrailed K52E, en la cual una interacción repulsiva (K17-K52) es sustituida por un puente salino (K17-E52), es 1,9 kcal/mol (11 °C en T_m) más estable que el polipéptido salvaje y 0,5 kcal/mol (0,7 °C en T_m) más estable que el mutante K52A (en el cual la interacción repulsiva está ausente) (Stollar *et al.*, 2003). Asimismo, se ha demostrado que en el homeodominio de vnd (NK-2), la sustitución H52R potencia la interacción electrostática entre la posición 52 y E17, incrementando la estabilidad térmica del polipéptido 7,7 °C a 0 M NaCl (Gonzalez *et al.*, 2001).

En conjunto, estos estudios sugieren que un puente salino conectando las posiciones 17 y 52 es bastante estabilizante. De hecho, una inspección cuidadosa de las secuencias de homeodominios disponibles permite comprobar que residuos hidrófobos en las posiciones 17 ó 52 aparecen con una frecuencia muy baja. Así, podemos encontrar residuos alifáticos (normalmente V o L) o aromáticos (Y en todos los casos) en la posición 17 en menos del 2% de las secuencias incluidas en nuestro conjunto de datos. Por lo que respecta a la posición 52, ésta no incorpora residuos alifáticos (V, L o I) en ningún caso, y sólo incorpora residuos aromáticos (F, W o Y) en un porcentaje menor del 0,5%. Por último, puentes hidrófobos (con residuos alifáticos o aromáticos indistintamente) conectando 17 y 52 no fueron encontrados en ninguna secuencia. Estas observaciones parecen sugerir que interacciones de este tipo en estas posiciones no son viables en homeodominios.

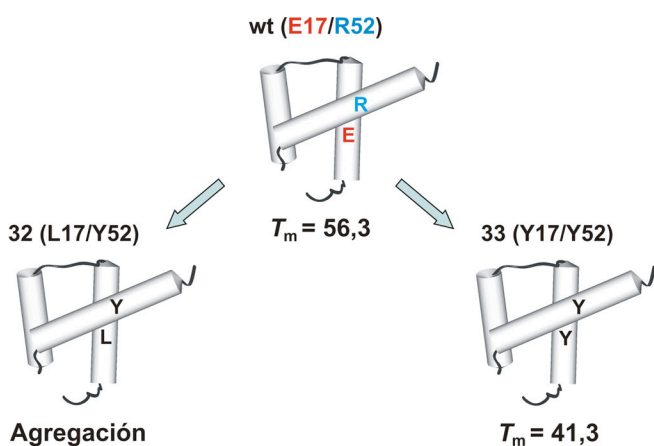


FIGURA 3.19: Representación esquemática del homeodominio de HESX-1 salvaje, junto con las variantes 32 y 33, con sustituciones en las posiciones 17-52. Para la variante 32, la estabilidad no pudo ser analizada debido a la marcada propensión a agregar y precipitar de la proteína. Se muestran los valores de T_m (en °C) medidos para la proteína salvaje y el doble mutante 33. Los detalles experimentales son los dados en la Tabla 3.1 y en la Figura 3.6.

Con objeto de verificar esta hipótesis se produjeron y analizaron las variantes **32** y **33**, con pares L17-Y52 e Y17-Y52, respectivamente. Una inspección de modelos moleculares sugiere que ambos pares de residuos serían tolerados estéricamente y podrían, en principio, aportar interacciones estabilizantes con el núcleo hidrófobo de la proteína. Los resultados obtenidos se representan esquemáticamente en la [Figura 3.19](#) y en la [Tabla 3.1](#).

Para la variante **33** (Y17/Y52), sólo se pudo purificar una cantidad mínima de proteína tras varios intentos, por lo que nuestro análisis se limita a la medida de una curva de desnaturalización térmica. No obstante, dicho experimento fue suficiente para verificar con claridad una caída significativa de estabilidad respecto a la proteína salvaje ($\Delta T_m = -15,0$ °C a 0 M NaCl). Por lo que respecta a la variante **32** (L17/Y52), esta proteína no pudo ser expresada soluble bajo las condiciones probadas, lo que parece indicar que la combinación L17/Y52 promueve la agregación y precipitación del polipéptido. En conjunto, estas observaciones sugieren que cadenas laterales alifáticas o aromáticas conectando las posiciones 17 y 52 reducen la estabilidad de la proteína y podrían también afectar a su solubilidad.

En conclusión, a diferencia de las preferencias observadas para las posiciones 19-30 y 31-42, sólo residuos polares parecen ser tolerados en las posiciones 17-52. Es posible que los residuos introducidos en las posiciones 17 y 52 sean incapaces de empaquetar correctamente con el núcleo hidrófobo de HESX-1. Adicionalmente, es preciso tener en cuenta que, a diferencia de 19-30 y 31-42, las posiciones 17-52 presentan un alto grado de exposición al disolvente, con lo que la incorporación de grupos hidrófobos en ellas podría promover la agregación del polipéptido. Este efecto podría verse potenciado por el carácter dinámico de las zonas mutadas (el residuo 52 está próximo a la hélice III-b, una región parcialmente desestructurada), lo cual podría incrementar aún más su exposición al disolvente y, por tanto, los problemas de solubilidad.

3.1.9. Influencia del empaquetamiento de las cadenas laterales en el efecto estabilizador de los puentes salinos

Además de sus aportaciones electrostáticas a la energética del plegamiento, las cadenas laterales cargadas también pueden contribuir a la estabilidad de las proteínas por medio de interacciones de empaquetamiento que involucren a sus grupos metileno. Estos contactos podrían ser especialmente importantes para los pares 19-30 y 31-42, debido a su particular localización en la estructura del homeodominio, ni totalmente expuestos al disolvente ni totalmente enterrados. De hecho, como ya hemos apuntando anteriormente, dichas

posiciones han sido clasificadas como fronterizas, ya que empaquetan contra el núcleo hidrófobo de la proteína, protegiéndolo de la exposición al disolvente.

Con el objetivo de analizar la estabilización derivada del empaquetamiento de los grupos $-\text{CH}_2-$ en estas posiciones, se prepararon dos variantes adicionales. En ambos casos, un ácido glutámico involucrado en una interacción de puente salino fue sustituido por un ácido aspártico, que incluye un metileno menos en la cadena lateral (ver [Figura 3.20](#)). De esta manera, el puente E19-R30 presente en la variante **19** fue reemplazado por el par D19-R30 para dar la variante **2** (D19/R30). De un modo similar, el residuo E42, involucrado en un puente salino con R31 en la proteína salvaje, fue sustituido por D, para dar la variante **27** (R31/D42) (ver [Tabla 3.1](#) y [Figura 3.20](#)).

En principio, el acortamiento de la cadena lateral ácida presente en los puentes salinos E19-R30 y R31-E42 podría tener un efecto pequeño en el balance electrostático de ambas interacciones. Sin embargo, dada la localización fronteriza del residuo 19, la ausencia de un solo grupo metileno podría conllevar la pérdida de interacciones de empaquetamiento importantes, siendo fuertemente desestabilizante. Por el contrario, la ausencia de este metileno en la posición 42, algo más expuesta al disolvente, debería tener un efecto mucho menos acusado en comparación. Los resultados obtenidos fueron plenamente consistentes con estos razonamientos.

De acuerdo con nuestro análisis, el puente salino D19-R30 presente en **2** (D19/R30) es fuertemente desestabilizante, no sólo con respecto al par hidrófobo presente en la proteína salvaje, sino también con respecto al puente salino E19-R30 presente en la variante **19** ($\Delta\Delta G_{\text{total}} = 1,64$ kcal/mol, ver [Tabla 3.1](#)). Esta marcada desestabilización puede resultar sorprendente dado el carácter aparentemente conservativo de la mutación E19D. Sin embargo, los datos obtenidos demuestran la importancia de los grupos $-\text{CH}_2-$ localizados en posiciones fronterizas. Puede apreciarse como el efecto desestabilizador de la mutación se mantiene en presencia de 2 M NaCl ($\Delta\Delta G_{\text{no polar}} = 1,69$ kcal/mol, ver [Figura 3.20](#)). Una fracción de esta diferencia, estimada en unos 0,2 kcal/mol, podría atribuirse a la menor propensión intrínseca de hélice α del ácido aspártico respecto al ácido glutámico (O'Neil y DeGrado, 1990; Chakrabartty *et al.*, 1994; Pace y Scholtz, 1998). Sin embargo, los 1,49 kcal/mol restantes apuntan a unas interacciones de empaquetamiento peor optimizadas en **2** (D19/R30), todo ello motivado por la ausencia de un único metileno. En cambio, los términos $\Delta G_{\text{coulomb}}$ son similares en **2** (D19/R30) y **19** (E19/R30), lo que indica que la interacción electrostática entre los residuos 19 y 30 no queda alterada significativamente por el acortamiento de la cadena lateral ácida.

En contraste con esta situación, y tal y como anticipábamos, la sustitución E→D en la posición 42 de la proteína salvaje, para dar la variante

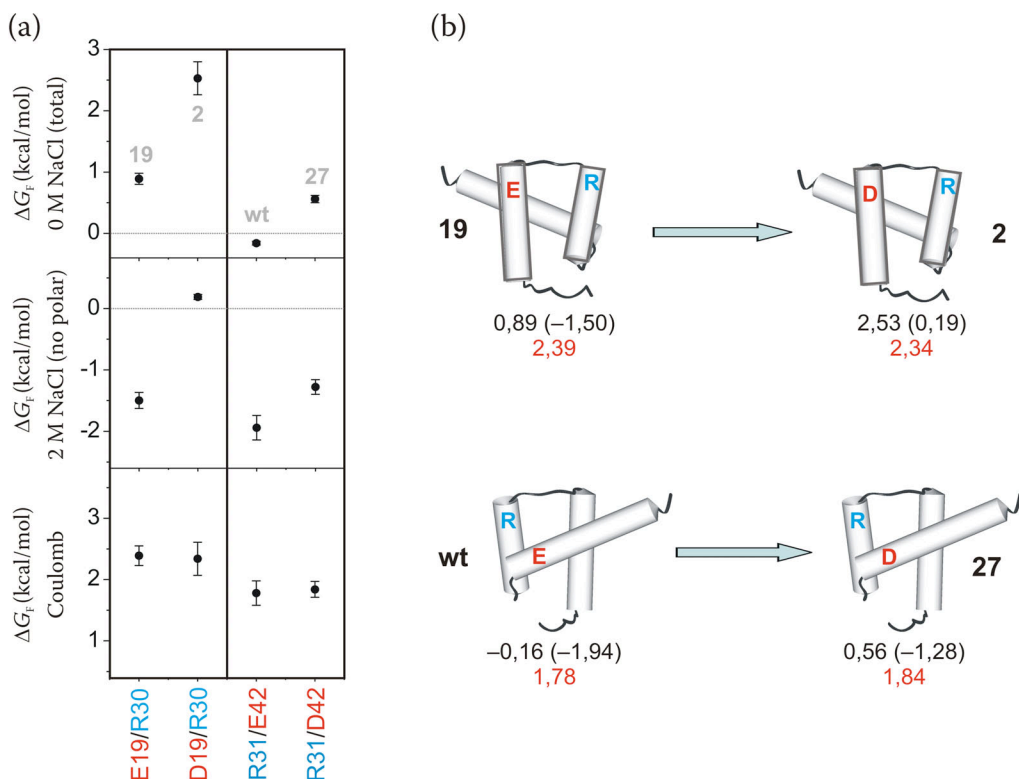


FIGURA 3.20: (a) Representación gráfica de los valores de ΔG_F a 55 °C obtenidos a partir de experimentos de desnaturalización térmica para las variantes del homeodominio de HESX-1 **19** y **wt**, así como para las variantes **2** y **27**, caracterizadas por sustituciones E→D en las posiciones 19 y 42, respectivamente. Se indican los valores de energía libre en presencia de 0 M NaCl (ΔG_{total} , panel superior) y 2 M NaCl ($\Delta G_{no\ polar}$, panel central) junto con la diferencia entre ambos ($\Delta G_{coulomb}$, panel inferior). (b) Representación esquemática de las variantes analizadas. Los residuos cargados negativamente se muestran en rojo, y los cargados positivamente en azul. Se dan los valores numéricos de ΔG_{total} , $\Delta G_{no\ polar}$ (entre paréntesis) y $\Delta G_{coulomb}$ (en rojo) a 55 °C obtenidos para cada variante a partir de experimentos de desnaturalización térmica. Los detalles experimentales se recogen en la [Tabla 3.1](#) y en la [Figura 3.6](#). Las energías libres se dan en kcal/mol.

27 (R31/D42), conduce a un descenso de estabilidad más moderado (de unos 0,7 kcal/mol tanto a 20 como a 55 °C, ver [Tabla 3.1](#)). Los motivos de este comportamiento vuelven a ser los mismos que los anteriormente expuestos. Así, puede observarse como las contribuciones electrostáticas a la estabilidad son casi idénticas en la proteína salvaje (R31/E42) y en la variante **27** (R31/D42). De acuerdo con los datos obtenidos, las contribuciones independientes de la sal son las responsables de la diferencia en estabilidad observada ($\Delta\Delta G_{no\ polar} = 0,66$ kcal/mol, ver [Figura 3.20](#)). Estos resultados

sugieren, una vez más, que la reducción del número de contactos hidrófobos y de van der Waals establecidos por la cadena lateral del ácido aspártico es el factor responsable de este efecto. Sin embargo, dada la localización algo más expuesta de la posición 42 en relación con la posición 19, la desestabilización observada en la variante **27** no es tan acusada como la anteriormente descrita para la variante **2**.

En conclusión, nuestros resultados indican que, si bien los puentes salinos E19-R30 y R31-E42 no parecen suponer un beneficio electrostático global, el empaquetamiento de los grupos metileno de sus cadenas laterales sí aporta unas significativas contribuciones favorables a ΔG_F . Esto podría tratarse de una característica general de los puentes salinos situados en posiciones fronterizas del plegamiento.

3.1.10. Análisis de valores pK_a en una selección de mutantes múltiples por RMN

Con el objetivo de analizar más en detalle la contribución electrostática de las cadenas de ácido aspártico y ácido glutámico a la estabilidad del homeodominio, llevamos a cabo la medición de los valores de pK_a en el estado plegado (pK_a^F) de todos estos residuos ácidos presentes en las variantes **12** (R19/E30/E15/R37) y **15** (E19/R30/R15/E37), mediante espectroscopía de RMN. Dichos homeodominios incluyen el mayor número potencial de puentes salinos en posiciones conservadas (R19-E30, E15-R37, R31-E42 y E17-R52 en **12** y E19-R30, R15-E37, R31-E42 y E17-R52 en **15**). En principio, es de esperar que interacciones electrostáticas favorables con el entorno se traduzcan en un decremento en el valor de pK_a^F con respecto al valor correspondiente al estado desnaturalizado (pK_a^U). Por el contrario, si la interacción electrostática del residuo ácido con el entorno es netamente repulsiva, es esperable un incremento en el pK_a^F . De esta forma, las variaciones de pK_a proporcionan un modo sencillo de determinar el efecto de una carga determinada sobre la estabilidad conformacional del polipéptido (Wyman, 1948; Yang y Honig, 1993; Marti *et al.*, 2000; Marti y Bosshard, 2003; Bosshard *et al.*, 2004). Con objeto de medir los incrementos de pK_a para cada residuo ácido, se monitorizaron las variaciones de desplazamiento químico con el pH de los protones adyacentes a los grupos carboxilo ($H\beta$ en cadenas laterales de ácido aspártico y $H\beta$ y $H\gamma$ en cadenas laterales de ácido glutámico). Dadas las dificultades de medir los pK_a^U de cada residuo ácido, se asumió un valor de 3,9 para el ácido aspártico y de 4,3 para el ácido glutámico (Nozaki y Tanford, 1967; Yang y Honig, 1993). Los resultados obtenidos se representan en la [Figura 3.21](#) y la [Tabla 3.4](#).

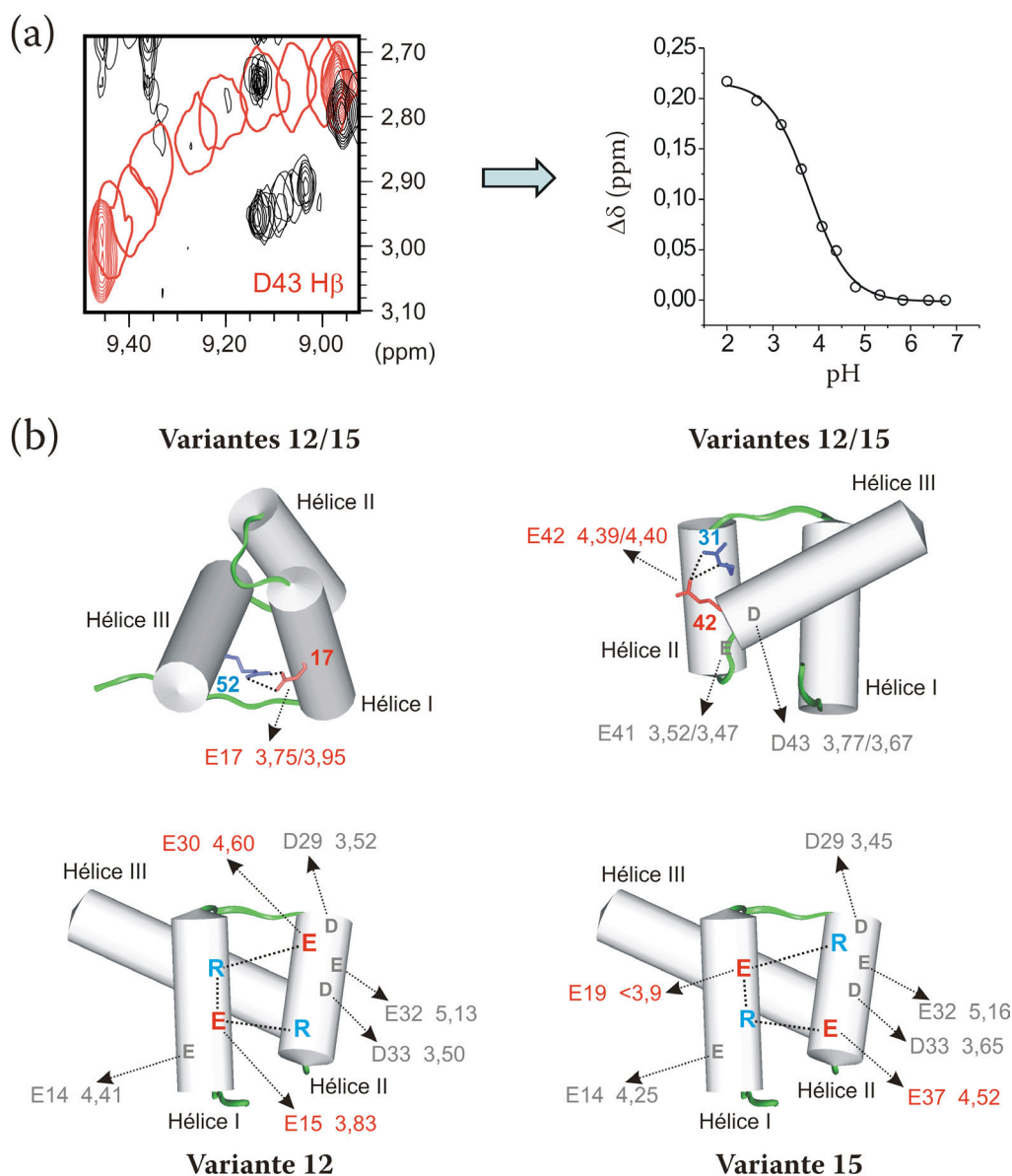


FIGURA 3.21: (a) Dependencia del desplazamiento químico de los protones H β 2/H β 3 con el pH en el cuádruple mutante **12**. La curva de valoración correspondiente se representa en el panel derecho. (b) Panel superior: En rojo se indican los valores de pK_a medidos para las cadenas laterales de E17 (izquierda) y E42 (derecha) en las variantes **12** y **15** (separados por una barra) empleando RMN. También se indican, en gris, los valores medidos para los residuos E41 y D43. Panel inferior: Valores de pK_a medidos para varias cadenas laterales de E y D en las variantes **12** (izquierda) y **15** (derecha) empleando RMN. Destacados en rojo se dan los valores correspondientes a los glutámicos incorporados por mutagénesis (posiciones 15 y 30 en **12**, y posiciones 19 y 37 en **15**). Los detalles experimentales se recogen en la [Tabla 3.4](#).

TABLA 3.4: Valores de pK_a^F medidos para las cadenas laterales de ácido aspártico y ácido glutámico en los mutantes **12** y **15**^a

<i>Residuo</i>	<i>Variante 12</i>		<i>Variante 15</i>	
	pK_a^F	$\Delta\Delta G_F$ (kcal/mol)	pK_a^F	$\Delta\Delta G_F$ (kcal/mol)
E14	4,41 ± 0,10	0,06 ± 0,06	4,25 ± 0,10	-0,03 ± 0,06
E15	3,83 ± 0,14	-0,26 ± 0,08	-	-
E17	3,75 ± 0,10	-0,30 ± 0,06	3,95 ± 0,11	-0,19 ± 0,06
E19	-	-	<3,9	<-0,22
D29	3,52 ± 0,08	-0,21 ± 0,04	3,45 ± 0,09	-0,25 ± 0,05
E30	4,60 ± 0,15	0,17 ± 0,08	-	-
E32	5,13 ± 0,12	0,46 ± 0,07	5,16 ± 0,13	0,47 ± 0,07
D33	3,50 ± 0,06	-0,22 ± 0,03	3,65 ± 0,07	-0,14 ± 0,04
E37	-	-	4,52 ± 0,10	0,12 ± 0,06
E41	3,52 ± 0,13	-0,43 ± 0,07	3,47 ± 0,11	-0,46 ± 0,06
E42	4,39 ± 0,10	0,05 ± 0,06	4,40 ± 0,10	0,06 ± 0,06
D43	3,77 ± 0,08	-0,07 ± 0,04	3,67 ± 0,10	-0,13 ± 0,06
E62	4,00 ± 0,11	-0,17 ± 0,03	4,10 ± 0,10	-0,11 ± 0,06
M67	3,40 ± 0,06	-	3,50 ± 0,08	-

^a Se muestran también los correspondientes valores $\Delta\Delta G_F$, calculados mediante la ecuación $\Delta\Delta G_F^i = 2,3RT(pK_a^{F,i} - pK_a^{U,i})$, considerando unos valores pK_a^U de 3,9 para ácido aspártico y 4,3 para ácido glutámico. Valores positivos de $\Delta\Delta G_F$ indican que la forma cargada del residuo es desestabilizante, y valores negativos que la forma cargada es estabilizante. Las valoraciones por RMN se llevaron a cabo a 5 °C en tampón fosfato 10 mM, 100 mM NaCl. Se indican también los errores estimados.

Los valores de los pK_a^F medidos para E17 en **12** y en **15** son de 3,75 y 3,95, respectivamente, lo que está de acuerdo con un efecto estabilizador moderado del residuo cargado, de entre 0,2 y 0,3 kcal/mol. Estos datos apoyan la contribución favorable del puente salino E17-R52 que hemos comentado. En cambio, la cadena lateral de E42 muestra un pK_a^F muy cercano al valor de 4,3 que hemos considerado para pK_a^U (4,39 y 4,40 para **12** y **15**, respectivamente), lo que sugiere que una carga negativa en esta posición supone una contribución mínima o nula a ΔG_F .

De acuerdo con nuestro análisis, un ácido glutámico cargado negativamente en la posición 30 del mutante **12** (R19/E30/E15/R37) tiene una influencia desfavorable en la estabilidad de la proteína (su pK_a^F es de 4,60). Probablemente la interacción atractiva entre R19 y E30 en **12** es contrarrestada por contactos repulsivos con otros residuos cercanos. De hecho, una inspección de la superficie del homeodominio en esa zona revela la presencia de tres residuos de carga negativa cerca de E30, que son D29, E32 y D33.

Por lo que respecta al residuo E15 en el mutante **12**, obtenemos un valor de pK_a^F de 3,83, lo que sugiere que la forma cargada de este residuo tiene un efecto claramente estabilizante, probablemente motivado por unas interacciones favorables con los residuos cercanos R19 y R37.

La superposición excesiva de las señales de protón de E19 en la variante **15** (E19/R30/R15/E37) imposibilitó una medida precisa del correspondiente pK_a^F . No obstante, pudimos establecer un límite superior de 3,9. Esto refleja la doble interacción electrostática favorable de E19 con R30 y R15.

Finalmente, la carga negativa del residuo 37 en la variante **15** parece tener un cierto efecto desfavorable sobre la estabilidad del homeodominio, a juzgar por el valor de pK_a^F medido, de 4,52. Una explicación plausible, que ya hemos comentado en nuestra aproximación mutacional, es que su contacto favorable con R15 es contrarrestado por una interacción repulsiva con D33.

Es de destacar que estas conclusiones están cualitativamente de acuerdo con aquellas derivadas de los estudios de desnaturalización térmica previamente descritos. En la [Figura 3.22](#) se muestra una comparación. Así pues, la influencia electrostática de una cadena lateral cargada negativamente en la estabilidad de la proteína puede ser estimada a partir del efecto sobre $\Delta G_{\text{coulomb}}$ de mutaciones neutralizantes. De este modo, la sustracción del residuo de carga negativa E15 en **12** mediante la mutación E15V, para dar la variante **13**, conduce a una contribución dependiente de la sal más desfavorable ($\Delta\Delta G_{\text{coulomb}} = 0,43$ kcal/mol), lo que está de acuerdo con el pK_a^F de 3,8 medido para este residuo ([Figura 3.22a](#)). En otro ejemplo, el valor de pK_a^F inferior a 3,9 estimado para E19 en la variante **15**, se corresponde con la caída de $\Delta G_{\text{coulomb}}$ asociada a la mutación E19V para dar la variante **5** ($\Delta\Delta G_{\text{coulomb}} = 0,67$ kcal/mol, [Figura 3.22b](#)). Conclusiones similares pueden derivarse para otros residuos ácidos analizados. Desafortunadamente, una comparación más cuantitativa entre los resultados obtenidos empleando ambas aproximaciones no es posible dado que no se conocen los pK_a en el estado desnaturalizado de HESX-1 (pK_a^U).

Como punto final, el análisis de los valores de pK_a del resto de residuos ácidos presentes en **12** y **15** revela la presencia de otras cargas cuyo efecto sobre la estabilidad conformacional del homeodominio es considerable. De este modo, el residuo E32 presenta un pK_a^F superior a 5 en las variantes **12** y **15**, lo que indica que su forma ionizada es claramente desestabilizante. Su localización entre cuatro residuos ácidos (D29, D33, E41 y E42) podría ser la causa de este efecto. Por el contrario, de acuerdo con los datos de pK_a^F obtenidos, la presencia de cargas negativas en D29 y D33 tiene un efecto netamente favorable sobre el plegamiento del homeodominio, tanto en **12** como en **15**, a pesar de las potenciales repulsiones electrostáticas de estos residuos entre sí y con otros cercanos (ver [Tabla 3.4](#)). Por tanto, parece lógico suponer que estos aspárticos deben establecer interacciones favorables adicionales con residuos básicos cercanos. Entre ellas, la más evidente es el puente salino intrahélice D33-R37 en la variante **12**, del que ya hemos hablado anteriormente, presente también en el homeodominio salvaje en la forma D33-K37 ([Figura 3.23](#)). Pese a su aparente papel estabilizante, puentes salinos de cualquier polaridad en estas

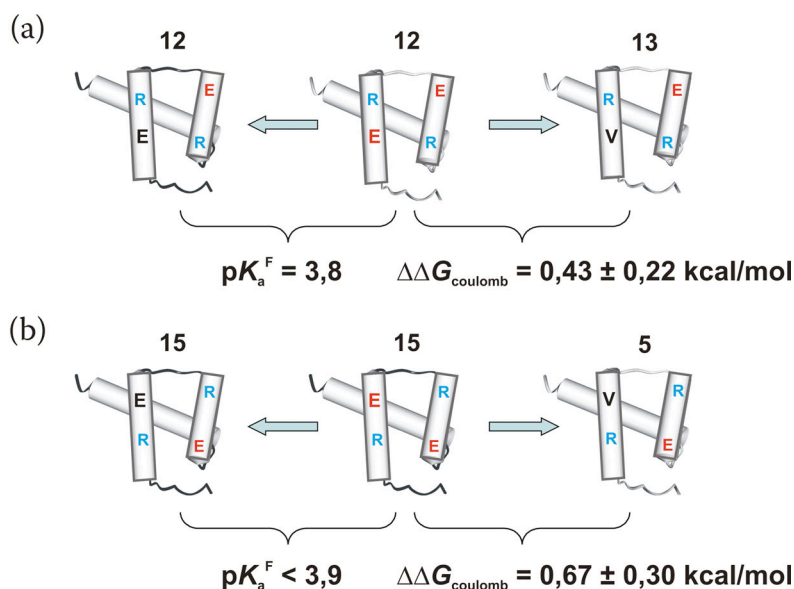
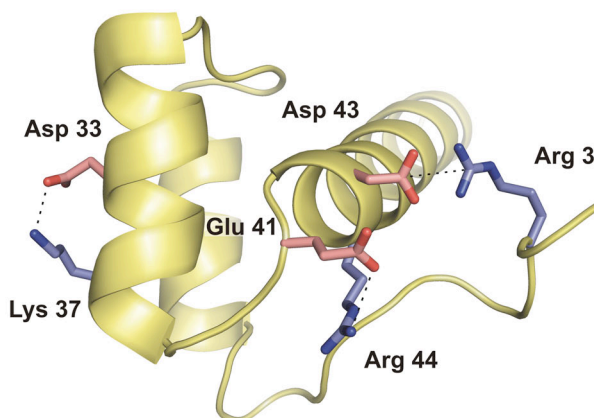


FIGURA 3.22: Efecto que la neutralización de las cargas negativas situadas en las posiciones (a) 15 y (b) 19 tiene sobre la estabilidad del homeodominio, determinado a partir de datos de pK_a^F , así como a partir del estudio termodinámico de mutantes. Se muestran los valores de pK_a^F , así como los términos $\Delta\Delta G_{\text{coulomb}}$ asociados a sustituciones puntuales de un ácido glutámico por un residuo neutro (ver “Procedimientos Experimentales” y [Tabla 3.4](#)). Puede comprobarse como los resultados obtenidos empleando ambas metodologías están cualitativamente de acuerdo. Los residuos con carga neutra se representan en negro, los cargados negativamente en rojo y los cargados positivamente en azul. Los detalles experimentales se recogen en las [Tablas 3.1](#) y [3.4](#).

posiciones pueden darse en tan solo el 14,7% de las 754 secuencias consideradas en nuestro análisis estadístico.

Otros ejemplos vienen dados por los residuos E41 y D43, cuyos valores de pK_a^F sugieren la existencia de otros puentes salinos adicionales en HESX-1 no conservados en la familia de los homeodominios ([Figura 3.23](#)). De este modo, E41 ($pK_a^F = 3,52$ en **12** y $pK_a^F = 3,47$ en **15**), perteneciente al giro II-III, parece contactar con R44, situado en el inicio de la hélice III. Puentes salinos de este tipo serían posibles en tan solo el 1,3% de los homeodominios. De mayor interés resulta el caso de D43 ($pK_a^F = 3,77/3,67$), cuya cadena lateral podría interaccionar electrostáticamente con R3, localizada en el brazo N-terminal del polipéptido. Como se ha explicado en la introducción, el brazo N-terminal tiene carácter dinámico y participa en el reconocimiento de ADN. Podría, por tanto, especularse que un posible papel funcional para este puente salino es contribuir a la unión a ADN orientando el brazo N-terminal y reduciendo su

FIGURA 3.23: Los datos de pK_a^F obtenidos, en combinación con la información estructural disponible, permite sugerir la presencia en HESX-1 de otros puentes salinos no conservados en homeodominios: D33-K37, E41-R44 y D43-R3.



entropía conformacional. Interacciones de este tipo serían factibles en tan solo el 7,3% de las secuencias de homeodominios.

En conclusión, el análisis de los valores de pK_a^F nos ha permitido arrojar más luz sobre las contribuciones electrostáticas de los residuos ácidos presentes en HESX-1. Éstos incluyen no sólo los involucrados en los puentes salinos conservados que hemos abordado en nuestra aproximación mutacional, sino otros que podrían tener un papel importante en la estabilidad del homeodominio de HESX-1.

3.1.11. Puentes salinos frente a pares hidrófobos como determinantes de la estabilidad del homeodominio: resumen de resultados e influencia del grado de exposición de los residuos involucrados

La exhaustiva comparación de mutantes del homeodominio de HESX-1 llevada a cabo, permite concluir que las interacciones hidrófobas aportan un efecto estabilizador superior al logrado por las interacciones electrostáticas cuando se localizan en posiciones fronterizas del polipéptido. La explicación principal debe buscarse en las contribuciones a la estabilidad independientes de la sal ($\Delta G_{\text{no polar}}$). Así, puede observarse como ΔG_{total} correlaciona en gran medida con $\Delta G_{\text{no polar}}$, pero mucho menos con $\Delta G_{\text{coulomb}}$, lo que indica que los cambios de estabilidad observados entre variantes son debidos principalmente a interacciones no electrostáticas, como pueden ser el efecto hidrófobo o las fuerzas de van der Waals. Esto queda ilustrado en las Figuras 3.24 y 3.25 (análogas a las Figuras 3.7 y 3.8, pero a las que se han añadido las variantes con mutaciones en las posiciones 26, 31 y 42).

Por lo que respecta a los pares de residuos situados en posiciones expuestas, tales como 17-52, nuestro análisis revela un comportamiento

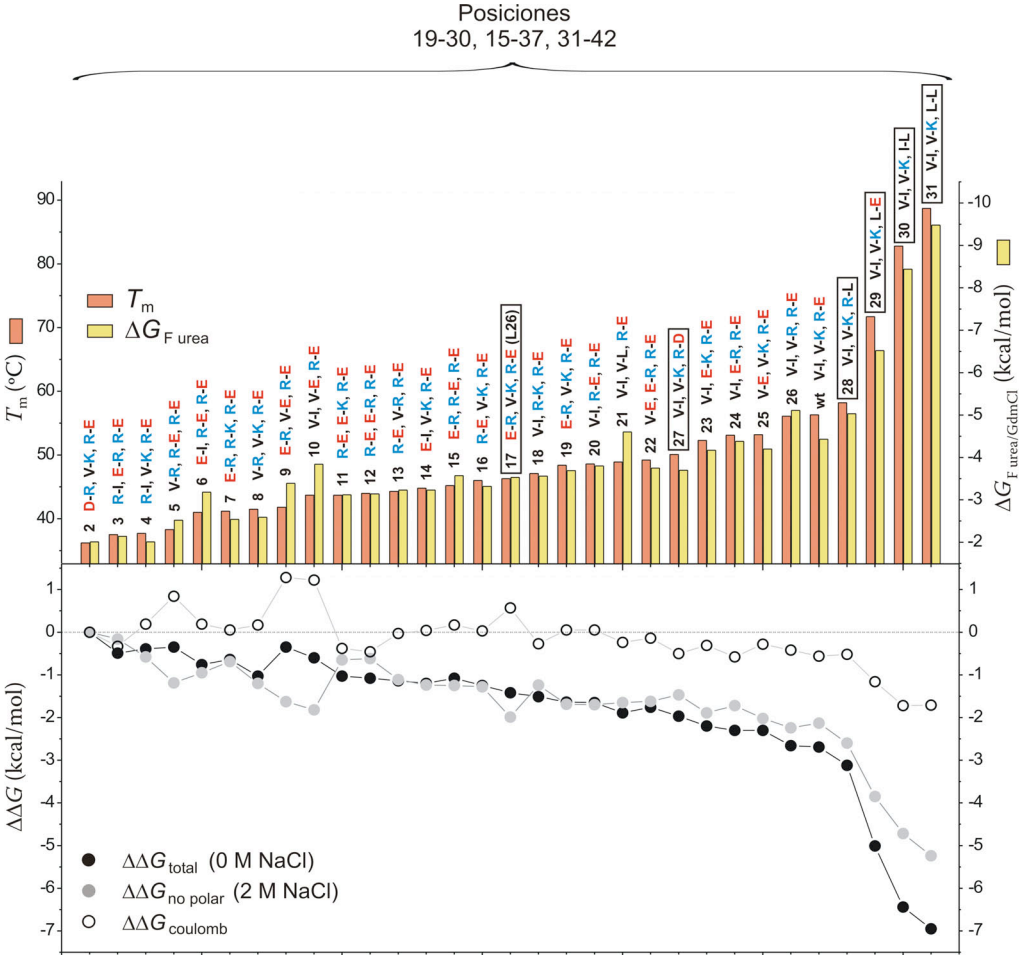


FIGURA 3.24: Representación gráfica de los parámetros termodinámicos medidos para todas las variantes analizadas en esta tesis. En el gráfico de columnas se representan, en rojo, las T_m medidas a 0 M NaCl en tampón fosfato 10 mM pH 6,0, 0,5 mM DTT, y en amarillo, los ΔG_F medidos mediante desnaturalización con urea (GdmCl para las variantes 28, 29, 30 y 31) a 20 °C, en tampón fosfato 10 mM pH 6,0, 0,5 mM DTT, 100 mM NaCl. En el panel inferior se representan las diferencias de energía libre a 55 °C respecto a la variante menos estable (variante 2), derivadas de los experimentos de desnaturalización térmica a 0 M NaCl ($\Delta\Delta G_{\text{total}}$), a 2 M NaCl ($\Delta\Delta G_{\text{no polar}}$), así como la diferencia entre ambos ($\Delta\Delta G_{\text{coulomb}}$). Valores negativos indican una energía libre más favorable que en la variante 2. Las variantes añadidas respecto a la Figura 3.7 se destacan con su etiqueta encuadrada. Para las variantes 28, 29, 30 y 31, los valores de $\Delta\Delta G$ se calcularon a partir de las diferencias con la proteína salvaje a 77 °C. Los errores no se representan para favorecer la claridad. Los residuos cargados negativamente se muestran en rojo, los cargados positivamente en azul, y los alifáticos en negro.

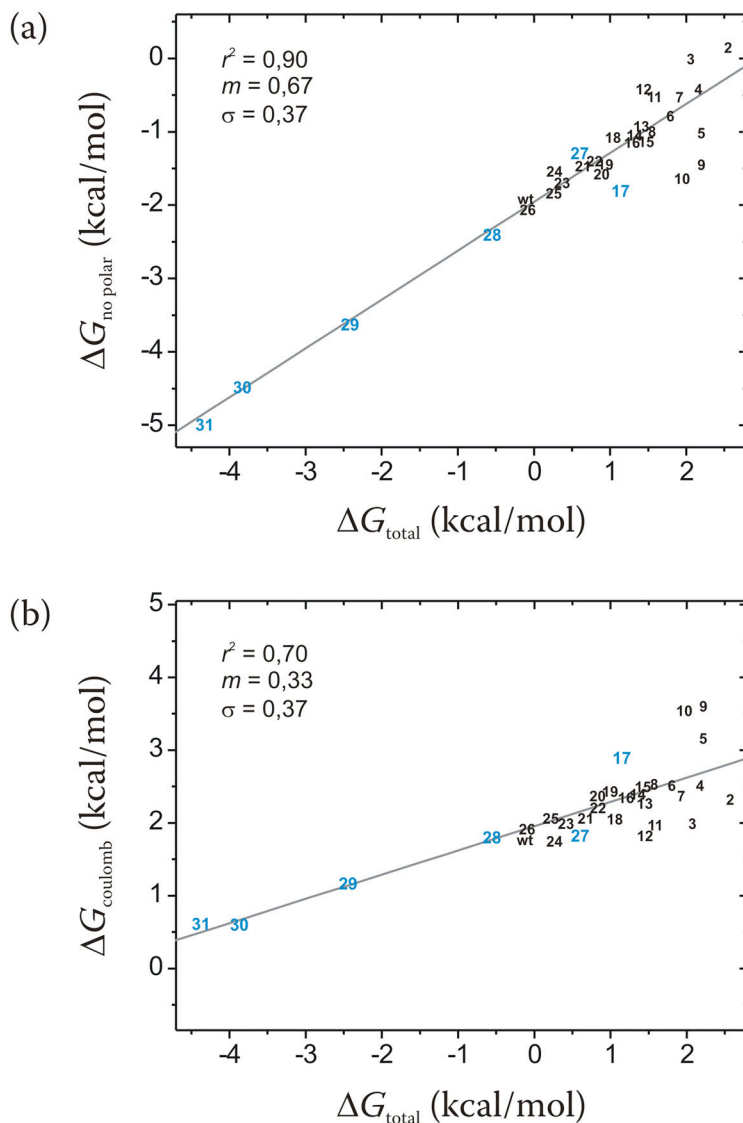


FIGURA 3.25: **Gráficas de correlación** entre ΔG_{total} y los términos (a) $\Delta G_{\text{no polar}}$ y (b) $\Delta G_{\text{coulomb}}$, a 55 °C, para todas las variantes del homeodominio analizadas en esta tesis. Los números en el interior de las gráficas identifican las variantes. En azul se muestran las que han sido incorporadas respecto a la Figura 3.8 (17, 27, 28, 29, 30 y 31). Para las variantes 28, 29, 30 y 31, los valores de ΔG se calcularon a partir de las diferencias con la proteína salvaje a 77 °C. En las dos gráficas y en ambos ejes se ha utilizado la misma escala para facilitar las comparaciones. Las líneas grises representan los ajustes de los datos por mínimos cuadrados. r^2 , coeficiente de correlación elevado al cuadrado. m , pendiente de la recta. σ , desviación típica.

opuesto al descrito para posiciones fronterizas. Así, hemos visto que la incorporación de residuos hidrófobos en estas posiciones, en las variantes **32** (L17/Y52) y **33** (Y17/Y52), da lugar a profundos problemas de estabilidad y solubilidad. Por lo tanto, parece claro que el puente salino E17-R52 (altamente conservado en homeodominios) no puede ser reemplazado por pares hidrófobos.

La conclusión final que podemos derivar de todo esto es que nuestros datos son consistentes con una preferencia de pares hidrófobos frente a puentes salinos en posiciones fronterizas, pero no así en posiciones más expuestas. La singular ubicación de los residuos fronterizos facilitaría la interacción favorable de sus cadenas laterales con el núcleo hidrófobo, lo que parece estar en el origen de los resultados obtenidos.

3.2

ANÁLISIS DE LAS PROPIEDADES DE UNIÓN A ADN DE UNA SELECCIÓN DE VARIANTES DEL HOMEODOMINIO

Como etapa final a nuestro trabajo, ensayamos las propiedades de asociación a ADN de una selección de variantes del homeodominio de HESX-1, utilizando microcalorimetría. Para ello se empleó un fragmento de ADN de 14 pb que incorpora la secuencia diana de HESX-1 (5'-TAATTG-3'). Las condiciones seleccionadas para nuestros experimentos (tampón fosfato 20 mM pH 6,0) fueron análogas a las previamente descritas en los estudios de estabilidad (ver parte experimental), con objeto de posibilitar una correlación directa entre unión a ADN y estabilidad. Adicionalmente, se incluyó en el tampón de calorimetría una concentración moderada de cloruro sódico e iones Mg^{2+} (150 mM NaCl, 5 mM $MgCl_2$) con la intención de limitar la asociación no específica y reducir la afinidad por ADN de las distintas variantes analizadas, incrementando, de este modo, la precisión en los valores de K_b obtenidos.

Los datos obtenidos se presentan en la [Tabla 3.5](#). De acuerdo con ellos, el homeodominio de HESX-1 salvaje presenta una afinidad por el ADN de $6,2 \cdot 10^6 \text{ M}^{-1}$, similar a la anteriormente obtenida por nuestro propio grupo de investigación, empleando medidas de dicróismo circular (de la Mata *et al.*, 2002). Puede apreciarse como, tal y como se esperaba considerando las condiciones experimentales seleccionadas, se trata de un valor de afinidad relativamente bajo en relación a lo habitualmente observado en homeodominios (en el rango de 10^6 a 10^{10} M^{-1}).

Un examen superficial de la [Tabla 3.5](#) revela que, tanto en la proteína salvaje como en todos los mutantes, la formación del complejo está impulsada entálpicamente, mientras que la entropía se opone a la unión. Este dato contrasta con los resultados descritos para la mayoría de los homeodominios en los que, como hemos apuntado en la introducción, la entropía suele ser ligeramente favorable. Como ejemplos, tenemos los casos de Antennapedia (Seimiya y Kurosawa, 1996; Dragan *et al.*, 2006), vnd (NK-2) (Gonzalez *et al.*, 2001), NKX2.5 (Fodor *et al.*, 2005) y TTF-1 (Del Vecchio *et al.*, 2008).

TABLA 3.5: Parámetros termodinámicos medidos para la unión de diferentes variantes del homeodominio de HESX-1 a un dúplex de ADN de 14 pb (5'-GTCTAATTGACGCG-3' y su hebra complementaria 5'-CGCGTCAATTAGAC-3') a 25 °C en tampón fosfato 20 mM pH 6,0, 150 mM NaCl, 5 mM MgCl₂, empleando microcalorimetría (ITC)

Variante	31-42	n^a	K_b^b (M ⁻¹)	ΔG_b^c (kcal/mol)	ΔH_b^d (kcal/mol)	$-T\Delta S_b^e$ (kcal/mol)
31	L-L	1,03	$1,3 \cdot 10^6$	$-8,3 \pm 0,2$	$-14,1 \pm 0,6$	$5,8 \pm 0,6$
29	L-E	1,02	$1,4 \cdot 10^6$	$-8,4 \pm 0,2$	$-14,9 \pm 0,5$	$6,5 \pm 0,5$
28	R-L	1,02	$8,1 \cdot 10^6$	$-9,4 \pm 0,2$	$-13,5 \pm 0,5$	$4,1 \pm 0,5$
wt	R-E	1,02	$6,2 \cdot 10^6$	$-9,2 \pm 0,1$	$-13,7 \pm 0,4$	$4,5 \pm 0,4$

Variante	19-30, 15-37	n^a	K_b^b (M ⁻¹)	ΔG_b^c (kcal/mol)	ΔH_b^d (kcal/mol)	$-T\Delta S_b^e$ (kcal/mol)
21	V-I, V-L	1,04	$6,5 \cdot 10^6$	$-9,3 \pm 0,1$	$-13,4 \pm 0,5$	$4,1 \pm 0,5$
wt	V-I, V-K	1,02	$6,2 \cdot 10^6$	$-9,2 \pm 0,1$	$-13,7 \pm 0,4$	$4,5 \pm 0,4$
2	D-R, V-K	1,01	$6,7 \cdot 10^6$	$-9,3 \pm 0,2$	$-15,3 \pm 0,6$	$6,0 \pm 0,6$
16	R-E, V-K	1,03	$5,8 \cdot 10^6$	$-9,2 \pm 0,1$	$-13,6 \pm 0,6$	$4,4 \pm 0,6$
17	E-R, V-K ^(L26)	1,01	$4,7 \cdot 10^6$	$-9,1 \pm 0,2$	$-14,1 \pm 0,5$	$5,0 \pm 0,5$
7	E-R, R-K	1,02	$5,5 \cdot 10^6$	$-9,2 \pm 0,1$	$-14,8 \pm 0,5$	$5,6 \pm 0,5$
11	R-E, E-K	1,02	$4,0 \cdot 10^6$	$-9,0 \pm 0,1$	$-14,1 \pm 0,4$	$5,1 \pm 0,4$
15	E-R, R-E	0,99	$5,3 \cdot 10^6$	$-9,1 \pm 0,1$	$-13,2 \pm 0,5$	$4,1 \pm 0,5$

^a n , número de sitios de unión idénticos e independientes de la proteína al ADN.

^b K_b , constante de afinidad de unión.

^c ΔG_b , variación de energía libre asociada a la unión.

^d ΔH_b , variación de entalpía asociada a la unión.

^e ΔS_b , variación de entropía asociada a la unión.

Tal y como se ha mencionado anteriormente, las posiciones 19, 30, 15 y 37 están localizadas lejos de la superficie de reconocimiento, mientras que el par 31-42 está directamente involucrado en el contacto con el ADN, de acuerdo con la información estructural disponible. Teniendo esto en cuenta, en la siguiente presentación de resultados se analizará por separado el efecto de las mutaciones en ambos grupos de posiciones sobre la asociación.

3.2.1. Propiedades de unión a ADN de las variantes con sustituciones en posiciones 31 y 42. Importancia del contacto R31-fosfato

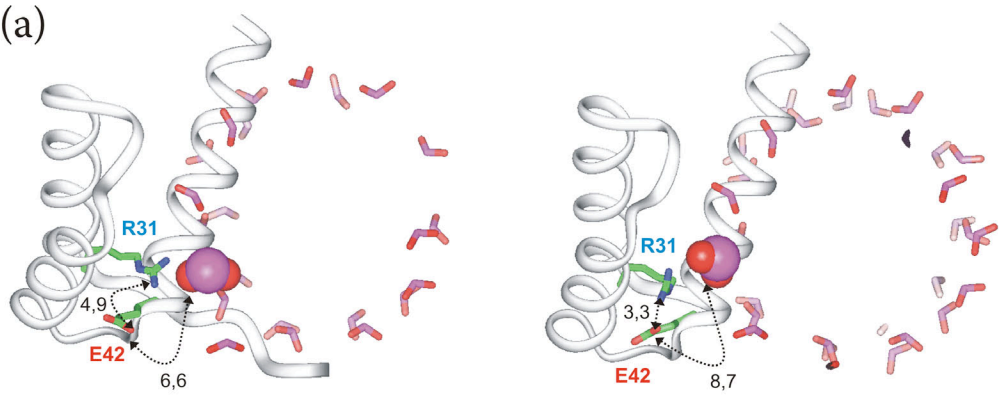
En primer lugar, analizamos la influencia del par 31-42 en los parámetros termodinámicos de la asociación homeodominio-ADN. Los datos estructurales disponibles indican que el residuo 31 participa en el proceso de reconocimiento, estableciendo contactos con un fosfato del esqueleto de ADN. Es interesante comprobar que, en la mayoría de los casos, esa interacción se establece sin necesidad de romper el puente salino con E42 preexistente. La

Figura 3.26a muestra la estructura de rayos X de los homeodominios de MSX-1 y Engrailed en su estado asociado. En ella, se destaca la posición de las cadenas R31 y E42. Puede observarse como, en ambos casos, R31 participa en una doble interacción de puente salino, que involucra la cadena lateral de E42 y un grupo fosfato de la doble hélice (resaltado en esferas en la **Figura 3.26a**). En principio, la influencia de este patrón de interacciones sobre la estabilidad del complejo resulta difícil de predecir. Por un lado, podría proponerse que el contacto R31-E42, presente en el polipéptido libre, tiene la misión de fijar la cadena lateral de R31 en su orientación correcta para la unión a ADN, preorganizando el receptor proteico y así reduciendo el coste entrópico del proceso de reconocimiento. Sin embargo, dicho efecto podría verse parcial o totalmente contrarrestado por interacciones repulsivas entre E42 y el esqueleto del ADN. A pesar de ello, una inspección de los datos estructurales disponibles revela que esta cadena lateral queda, en la mayoría de los casos, lejos de los fosfatos (de 6,5 a 9,0 Å, ver **Figura 3.26a**) y, por tanto, en condiciones fisiológicas, la repulsión electrostática entre ambos grupos no debería ser muy grande.

Las valoraciones por ITC llevadas a cabo con el homeodominio de HESX-1 salvaje (R31/E42) y con las variantes **28** (R31/L42), **29** (L31/E42) y **31** (L31/L42) a 25 °C se representan en la **Figura 3.26b**. Por su parte, los parámetros termodinámicos derivados de ellas se recogen en la **Tabla 3.5**. De acuerdo con estos datos, la mutación E42L (variante **28**) tiene una influencia prácticamente despreciable en las propiedades de asociación del polipéptido, indicando que la repulsión electrostática anteriormente postulada entre E42 y los fosfatos del ADN no es muy acusada, o bien que su efecto es compensado por la preorganización de R31 en la proteína salvaje.

Por el contrario, la mutación simple R31L tiene un efecto detectable sobre la estabilidad del complejo. Así, puede apreciarse como la variante **29** presenta una constante de afinidad (K_b) 4,4 veces inferior a la de la proteína salvaje, lo que corresponde a una disminución de energía libre de unión ($\Delta\Delta G_b$) de 0,9 kcal/mol (ver **Tabla 3.5**). Esta desestabilización probablemente refleja la ausencia, en dicha variante, de la interacción electrostática entre R31 y el fosfato del ADN. Cabe destacar que, a pesar de todo, el efecto observado es relativamente modesto. De hecho, la supresión, mediante mutagénesis, de otras interacciones arginina-fosfato similares en homeodominios ha llevado, en ocasiones, a desestabilizaciones drásticas del complejo proteína/ADN. A modo de ejemplo, tal y como hemos apuntado en la sección de antecedentes, estudios previos, realizados por nuestro grupo de trabajo, han mostrado que, en HESX-1, la sustitución simple de R53 (arginina muy conservada, también involucrada en contactos con el esqueleto del ADN [ver la introducción]) por cisteína (R53C), reduce la afinidad de unión más de tres órdenes de magnitud (de la Mata *et al.*, 2002). Otro ejemplo, si bien menos claro, es la mutación

(a)



(b)

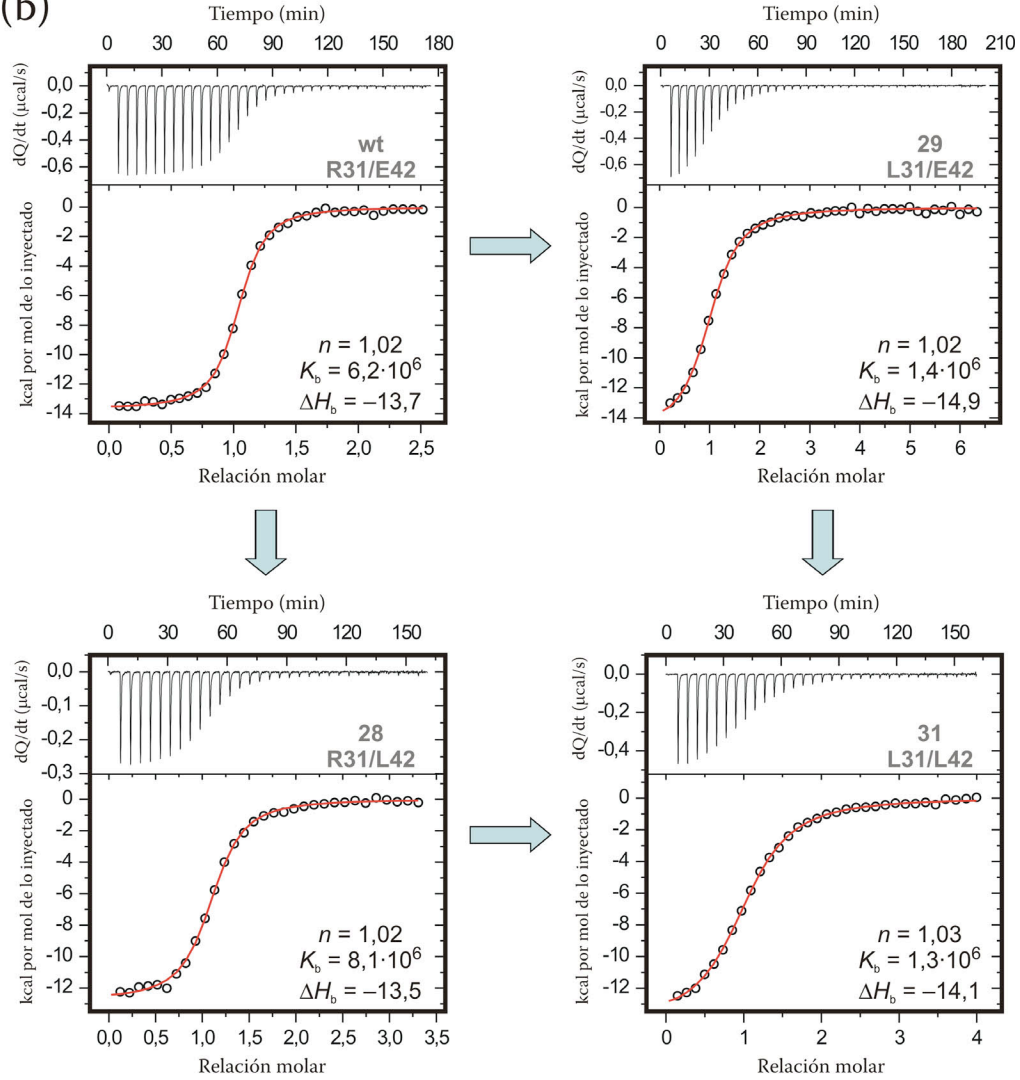


FIGURA 3.26 (PÁGINA ANTERIOR): (a) Representación de los contactos proteína-ADN en que está involucrado el par R31-E42 en los homeodominios de MSX-1 (izquierda, código PDB: 1IG7) y Engrailed (derecha, código PDB: 1DU0), de acuerdo con estudios de difracción de rayos X. En ambos casos se muestra en Å la distancia menor entre la cadena lateral del glutámico y los fosfatos del ADN. (b) Perfiles de ITC obtenidos mediante la valoración de disoluciones de ADN dúplex (5'-GTCTAATTGACGCG-3' y su hebra complementaria 5'-CGCGTCAATTAGAC-3') con el homeodominio de HESX-1 salvaje así como con las variantes **28**, **29** y **31** (con sustituciones simples o dobles en las posiciones 31-42). Dichos experimentos se llevaron a cabo a 25 °C en tampón fosfato 20 mM pH 6,0, 150 mM NaCl, 5 mM MgCl₂. En cada perfil, el panel superior muestra los datos experimentales sin procesar correspondientes a las inyecciones de proteína en la solución de ADN. La señal integrada, tras restar el calor de dilución, da el calor por mol de proteína inyectada, que se representa respecto a la relación molar proteína/ADN en el panel inferior de cada perfil. La línea roja representa el ajuste de mínimos cuadrados de los datos, asumiendo un modelo de unión 1:1. Se muestran también los parámetros termodinámicos derivados (véase también la [Tabla 3.5](#)). K_b se da en M⁻¹ y ΔH_b en kcal/mol.

R53H en Proboscipedia, que conlleva una caída de 20 veces en K_b (Benassayag *et al.*, 1997). Estas observaciones demuestran claramente que el papel de las interacciones arginina-fosfato en el reconocimiento de ADN es tremendamente dependiente de contexto.

Como punto final, puede comprobarse el carácter aditivo de las sustituciones sobre la estabilidad del complejo: el efecto de la doble sustitución es prácticamente idéntico a la suma de los efectos de las mutaciones simples (ver [Figura 3.26b](#)). De acuerdo con esto, la interacción electrostática entre las cadenas laterales de R31 y E42 no ejerce una influencia significativa sobre la afinidad de la asociación. Por lo tanto, el papel de E42 en la preorganización de R31, anteriormente propuesto, podría desestimarse.

Como conclusión, la cadena lateral de R31 está involucrada en el reconocimiento de ADN, tal como sugiere la información estructural disponible, si bien su papel es modesto en comparación con el de otras argininas del propio homeodominio. Por lo que respecta a E42, nuestros datos descartan su participación en el complejo, bien directamente o bien a través de su interacción electrostática con R31. Estos resultados contrastan con el efecto que, como hemos visto, tienen ambas posiciones en la estabilidad conformacional del homeodominio de HESX-1.

3.2.2. Propiedades de unión a ADN de las variantes con sustituciones en las posiciones 19, 30, 15 y 37. Relación estabilidad-función en el homeodominio

Finalmente, analizamos el efecto que mutaciones en las posiciones 19-30 y 15-37 tienen sobre la unión a ADN. Es preciso destacar que, al contrario que el

par R31-E42, los residuos 19, 30, 15 y 37 se localizan lejos de la interfase de interacción proteína-ADN (como se ilustra en la [Figura 3.27a](#)), y por lo tanto no participan en contactos con éste último. Sin embargo, tal como planteábamos en la sección de objetivos 2.2, podrían afectar indirectamente en la unión a ADN, modulando las propiedades dinámicas del polipéptido y, por tanto, su adaptabilidad. Esta hipótesis se ve reforzada por el hecho de que dichas posiciones tienen un efecto significativo sobre la estabilidad conformacional del

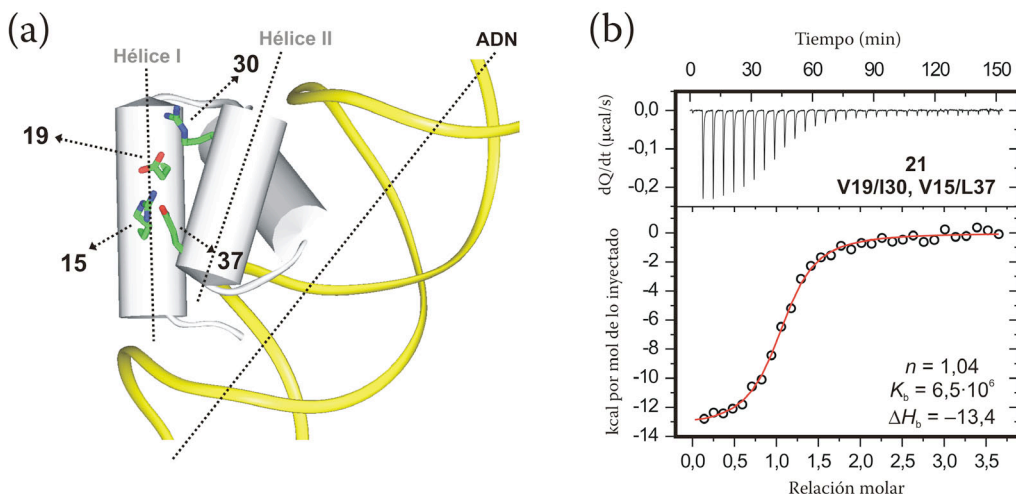
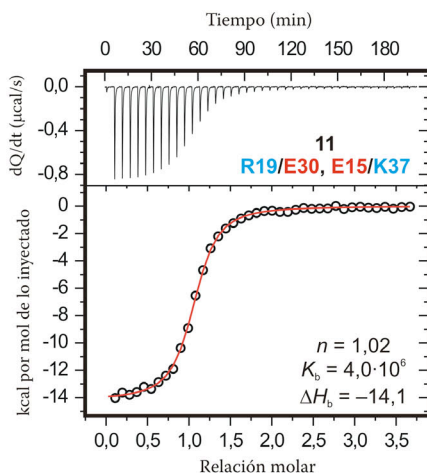


FIGURA 3.27: (a) Representación de las cadenas laterales en posiciones 19-30 y 15-37 presentes en el complejo formado por el homeodominio de Engrailed y el ADN, de acuerdo con datos de rayos X (código PDB: 1DU0). Los residuos que ocupan estas posiciones en HESX-1 salvaje son V19-I30 y V15-K37. (b) Perfiles de ITC obtenidos mediante la valoración de disoluciones de ADN dúplex (5'-GTCTAATTGACGCG-3' y su hebra complementaria 5'-CGCGTCAATTAGAC-3') con las variantes de HESX-1 **21** y **11**, que incorporan tétradas alifáticas o cargadas, respectivamente, en posiciones 19-30 y 15-37. Los experimentos se realizaron a 25 °C en tampón fosfato 20 mM pH 6,0, 150 mM NaCl, 5 mM MgCl₂. Al igual que en la [Figura 3.26](#), en cada perfil el panel superior muestra los datos experimentales sin procesar, el inferior la señal integrada, y la línea roja el ajuste de los datos asumiendo un modelo de unión 1:1. Se muestran también los parámetros termodinámicos derivados (véase también la [Tabla 3.5](#)). K_b se da en M^{-1} y ΔH_b en kcal/mol.



homeodominio, tal y como hemos demostrado. Las variantes **2** a **26**, descritas en la presente memoria, con diferencias máximas de ΔG_F de hasta 3,1 kcal/mol, proporcionan una excelente oportunidad para analizar la relación entre estabilidad y función en esta importante familia de proteínas.

La [Tabla 3.5](#) recoge los parámetros termodinámicos obtenidos para una selección de mutantes, con diferentes combinaciones de residuos cargados o hidrófobos en las posiciones 19-30 y 15-37. Puede apreciarse como, sorprendentemente, dichos datos no avalan la existencia de correlación alguna entre la estabilidad del polipéptido y la del complejo proteína/ADN. De hecho, las máximas diferencias en energía libre de unión ($\Delta\Delta G_b$) entre las proteínas ensayadas no superan las 0,3 kcal/mol. Bajo las condiciones experimentales empleadas, ni la estabilidad de la proteína (las máximas diferencias en ΔG_F , medidas en experimentos de desnaturalización con urea, para los mutantes seleccionados alcanzan los 2,6 kcal/mol, lo que equivale a 20 °C en T_m , ver [Tabla 3.1](#)) ni su carga total (que varía de -1 a +1 con respecto al homeodominio salvaje) parecen tener un efecto significativo en la unión a ADN. En la [Figura 3.27b](#) se puede comprobar una clara similitud entre los perfiles de ITC de las variantes **21** (V19/I30/V15/L37) y **11** (R19/E30/E15/K37), a pesar de que los residuos que incorpora cada una en las posiciones 19, 30, 15 y 37 son de naturaleza opuesta (alifáticos y cargados, respectivamente).

En la [Figura 3.28a](#) se representa la dependencia entre la estabilidad de las variantes analizadas y la estabilidad de sus complejos con ADN. Comparando esta gráfica con las de la [Figura 2.2](#) de la parte de antecedentes y objetivos, podemos comprobar fácilmente que el comportamiento del homeodominio de HESX-1 se corresponde con el caso (d) allí descrito, es decir, ausencia de correlación entre ambos parámetros, y poca dispersión de datos. Sin embargo, en nuestro caso esta dispersión es incluso menor que la mostrada en el ejemplo de la [Figura 2.2d](#), lo que apunta a que, efectivamente, las mutaciones en las posiciones 19, 30, 15 y 37 tienen una marcada influencia en la estabilidad del homeodominio sin prácticamente afectar a su afinidad por ADN.

Una inspección de los valores de ΔH_b y $T\Delta S_b$ obtenidos proporciona pistas relativas al origen de este comportamiento. Puede apreciarse como ambos parámetros correlacionan en cierta medida con la estabilidad de las variantes ([Figura 3.28b](#)). Así, las variantes menos estables tienden a presentar contribuciones entálpicas más favorables y contribuciones entrópicas menos favorables. De acuerdo con estos datos, la ausencia de correlación observada entre estabilidad y afinidad parece ser la consecuencia de un fenómeno de compensación entalpía-entropía (Dunitz, 1995; Cooper *et al.*, 2001).

Tal y como se señalaba en la introducción, diversos estudios han mostrado que la interacción específica de homeodominios con ADN viene acompañada del plegamiento de las regiones desordenadas N- y C-terminal del

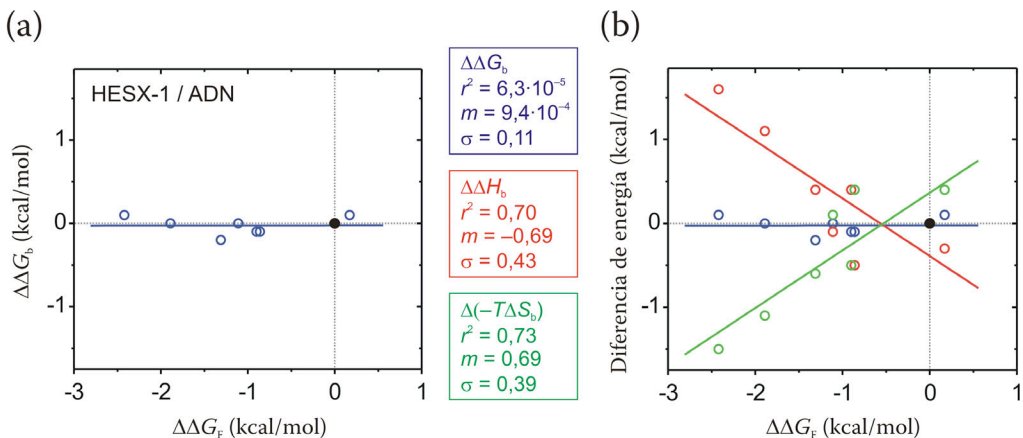


FIGURA 3.28: (a) Gráfica de correlación estabilidad-función para una selección de variantes del homeodominio de HESX-1 con distintas combinaciones de puentes salinos o hidrófobos en las posiciones 19, 30, 15 y 37. (b) La misma gráfica que en (a), pero a la que se han añadido los correspondientes datos de $\Delta\Delta H_b$ (rojo) y $\Delta(-T\Delta S_b)$ (verde). En ambas representaciones, se utiliza la misma escala para los dos ejes con objeto de facilitar las comparaciones. El punto correspondiente a la proteína salvaje se representa en negro y relleno. Los valores $\Delta\Delta G_F$ se calcularon a partir de los ΔG_F medidos a 20 °C mediante experimentos de desnaturalización con urea ($\Delta\Delta G_F = \Delta G_F^{\text{salvaje}} - \Delta G_F^{\text{mutante}}$). Los valores negativos indican menor estabilidad que la proteína salvaje, y los positivos mayor estabilidad. Asimismo, $\Delta\Delta G_b = \Delta G_b^{\text{salvaje}} - \Delta G_b^{\text{mutante}}$, $\Delta\Delta H_b = \Delta H_b^{\text{salvaje}} - \Delta H_b^{\text{mutante}}$, y $\Delta(-T\Delta S_b) = (-T\Delta S_b^{\text{salvaje}}) - (-T\Delta S_b^{\text{mutante}})$, con lo que, en los tres casos, los valores negativos indican una energía de unión menos favorable que para la proteína salvaje, y los positivos una energía más favorable. Los errores no se representan para favorecer la claridad. Los detalles experimentales se recogen en las Tablas 3.1 y 3.5. Las líneas azul, roja y verde representan el ajuste por mínimos cuadrados de los datos correspondientes a $\Delta\Delta G_b$, $\Delta\Delta H_b$ y $\Delta(-T\Delta S_b)$, respectivamente. r^2 , coeficiente de correlación elevado al cuadrado. m , pendiente de la recta. σ , desviación típica.

polipéptido. Dicha pérdida de grados de libertad conformacional podría ser extensiva al núcleo estructurado del homeodominio, cuyo carácter altamente dinámico en el estado libre ha sido puesto de manifiesto por estudios previos. De hecho, se ha sugerido que el puente salino 19-30 es parte de una red de covarianza de residuos que juega un papel en el ajuste inducido de la proteína en la formación del complejo (Stollar *et al.*, 2003). Con objeto de profundizar en esta dirección, se midió la dependencia de los parámetros termodinámicos con la temperatura. Para ello, llevamos a cabo, con un conjunto reducido de mutantes, valoraciones de ITC a tres temperaturas diferentes (15, 25 y 35 °C) tal y como se muestra en la Tabla 3.6 y en las Figuras 3.29 y 3.30. Estos experimentos nos permitieron estimar los valores de $\Delta C_{p,b}$ asociados a la formación del complejo, los cuales proporcionaron información sobre la

TABLA 3.6: Parámetros termodinámicos medidos para la unión del homeodominio de HESX-1 salvaje y las variantes **7**, **11** y **16** a un dúplex de ADN de 14 pb (5'-GTCTAATTGACGCG-3' y su hebra complementaria 5'-CGCGTCAATTAGAC-3') a 15, 25 y 35 °C en tampón fosfato 20 mM pH 6,0, 150 mM NaCl, 5 mM MgCl₂, empleando microcalorimetría (ITC)

Variante	19-30, 15-37	T (°C)	n ^a	K _b ^b (M ⁻¹)	ΔG _b ^c (kcal/mol)	ΔH _b ^d (kcal/mol)	-TΔS _b ^e (kcal/mol)	ΔC _{p,b} ^f (kcal/[mol·K])
wt	V-I, V-K	15	1,04	1,7·10 ⁷	-9,8 ± 0,1	-10,3 ± 0,3	0,5 ± 0,3	-0,38 ± 0,03
		25	1,02	6,2·10 ⁶	-9,2 ± 0,1	-13,7 ± 0,4	4,5 ± 0,4	
		35	0,99	2,6·10 ⁶	-8,7 ± 0,2	-18,0 ± 0,6	9,3 ± 0,6	
16	R-E, V-K	15	0,97	1,6·10 ⁷	-9,8 ± 0,1	-10,0 ± 0,3	0,2 ± 0,3	-0,47 ± 0,06
		25	1,03	5,8·10 ⁶	-9,2 ± 0,1	-13,6 ± 0,6	4,4 ± 0,6	
		35	1,01	3,0·10 ⁶	-8,8 ± 0,2	-19,7 ± 0,6	10,9 ± 0,6	
11	R-E, E-K	15	1,02	1,2·10 ⁷	-9,6 ± 0,1	-10,6 ± 0,3	1,0 ± 0,3	-0,41 ± 0,04
		25	1,02	4,0·10 ⁶	-9,0 ± 0,1	-14,1 ± 0,4	5,1 ± 0,4	
		35	1,02	2,0·10 ⁶	-8,6 ± 0,2	-19,3 ± 0,6	10,7 ± 0,6	
7	E-R, R-K	15	0,99	1,3·10 ⁷	-9,7 ± 0,1	-10,5 ± 0,3	0,8 ± 0,3	-0,48 ± 0,05
		25	1,02	5,5·10 ⁶	-9,2 ± 0,1	-14,8 ± 0,5	5,6 ± 0,5	
		35	0,99	2,7·10 ⁶	-8,8 ± 0,2	-20,7 ± 0,8	11,9 ± 0,8	

^a n, número de sitios de unión idénticos e independientes de la proteína al ADN.

^b K_b, constante de afinidad de unión.

^c ΔG_b, variación de energía libre asociada a la unión.

^d ΔH_b, variación de entalpía asociada a la unión.

^e ΔS_b, variación de entropía asociada a la unión.

^f ΔC_{p,b}, diferencia de capacidad calorífica asociada a la unión.

magnitud del ajuste inducido que acompaña a la asociación en cada caso (Spolar y Record, 1994).

Los resultados obtenidos muestran como, en el caso de la proteína salvaje, la entalpía asociada al proceso de unión aumenta con la temperatura (Tabla 3.6 y Figuras 3.29 y 3.30). La variación medida es consistente con un valor de ΔC_{p,b} negativo, de -0,38 kcal/(mol·K), en consonancia con lo descrito en otros homeodominios (ver introducción). Para las variantes **16**, **11** y **7**, que incluyen puentes salinos entre 19 y 30 de diferentes polaridades, se dedujo un comportamiento similar. Nuestros datos experimentales sugieren un pequeño incremento de ΔC_{p,b} para estas variantes, que podría estar relacionado con sus estabilidades térmicas. Esto sería consistente con un ajuste asociado a la unión un poco más pronunciado para las variantes más inestables. No obstante, las diferencias observadas no superan, en ninguno de los casos, las 0,1 kcal/(mol·K), y por lo tanto están próximas al error experimental de la medida. En definitiva, no se observan variaciones importantes de ΔC_{p,b} para las variantes menos estables, lo que parece sugerir que la mayor flexibilidad de las mismas no se traduce en un ajuste inducido más pronunciado.

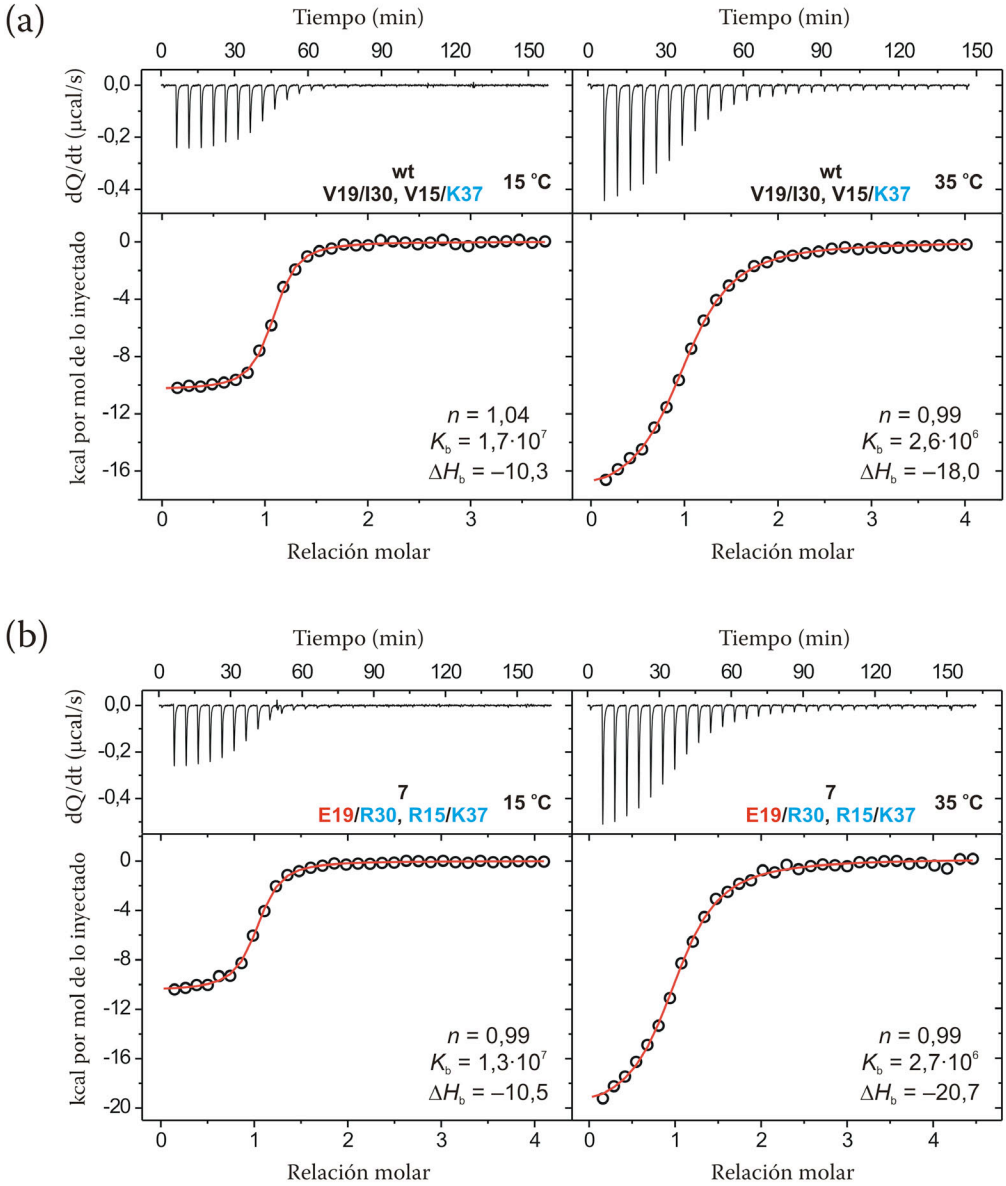


FIGURA 3.29: **Perfiles de ITC** obtenidos mediante la valoración de disoluciones de ADN dúplex (5'-GTCTAATTGACGCG-3' y su hebra complementaria 5'-CGCGTCAATTAGAC-3') con el homeodominio de HESX-1 (a) salvaje y (b) la variante 7, a 15 °C (izquierda) y 35 °C (derecha), en tampón fosfato 20 mM pH 6,0, 150 mM NaCl, 5 mM MgCl₂. En cada perfil, el panel superior muestra los datos experimentales sin procesar, el inferior la señal integrada, y la línea roja el ajuste de los datos asumiendo un modelo de unión 1:1. El incremento de ΔH_b con la temperatura resulta evidente y es consistente con valores de $\Delta C_{p,b}$ negativos. Se muestran también los parámetros termodinámicos derivados (véase también la [Tabla 3.6](#)). K_b se da en M^{-1} y ΔH_b en kcal/mol.

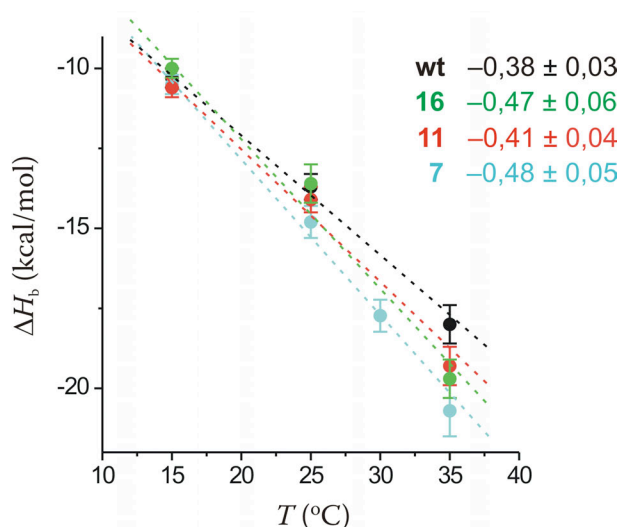


FIGURA 3.30: **Variaciones de entalpía para la unión específica de una selección de variantes del homeodominio de HESX-1 al ADN en función de la temperatura.** Los detalles experimentales son los dados en la [Tabla 3.6](#). Las variaciones observadas son consistentes con valores de $\Delta C_{p,b}$ negativos. Cada conjunto de datos se ha ajustado a una recta por mínimos cuadrados, y se muestra la pendiente de cada recta ($\Delta C_{p,b}$ de cada variante, en kcal/[mol·K]).

En conjunto, los resultados obtenidos descartan un efecto importante de las posiciones 19, 30, 15 y 37 del homeodominio en el reconocimiento de ADN. Además, indican que el potencial ajuste conformacional de las regiones estructuradas del polipéptido tiene un efecto mínimo sobre la afinidad de la asociación.

Como punto secundario, nuestros datos parecen descartar también la participación de la posición 26 y su red de interacciones en el proceso de asociación, tal y como ha sido propuesto en base a estudios previos (Clarke, 1995; Stollar *et al.*, 2003). De hecho, la [Tabla 3.5](#) muestra que los parámetros de unión a ADN de la variante 17 (E19/L26/R30) son muy parecidos a los del resto de mutantes (todos ellos con prolina en posición 26).

4 DISCUSIÓN

Los homeodominios son los motivos de unión a ADN que tienen más presencia en el entramado de factores de transcripción encargado de regular el desarrollo embrionario en eucariotas. Su evidente relevancia biológica ha despertado el interés de numerosos grupos de investigación, cuyo trabajo ha permitido una descripción detallada de esta familia de proteínas. En relación a su estudio, es preciso destacar que los homeodominios reúnen cualidades interesantes, tales como un asombroso grado de conservación de secuencia y estructura, o una plena funcionalidad aun cuando se encuentran aislados del resto de la homeoproteína. Estas características han facilitado su análisis en una amplia variedad de especies, posibilitando la obtención de valiosos datos comparativos y conclusiones de calado evolutivo.

El trabajo desarrollado en esta tesis aporta nuevos datos sobre las relaciones entre secuencia, estructura, estabilidad y función en homeodominios. El sistema que hemos seleccionado como objeto de estudio es el motivo de reconocimiento de ADN del factor de transcripción humano HESX-1. Dicho polipéptido se caracteriza por ser el homeodominio natural más estable descrito hasta la fecha, con una T_m de 56 °C en ausencia de sal. No obstante, dicha estabilidad no puede considerarse elevada si tenemos en cuenta que debe ejercer su función a 37 °C. De hecho, tal y como hemos mencionado anteriormente, una de las propiedades más frecuentes en homeodominios es su estabilidad marginal. Esta característica les confiere tiempos de vida cortos en el interior celular, garantizando que su acción no se prolongue descontroladamente en el tiempo. Por otro lado, la función biológica de estas proteínas depende críticamente de su capacidad de reconocer con elevada afinidad su secuencia diana de ADN. Parece claro que, dada su relevancia, ambas propiedades, estabilidad y función, han debido estar finamente ajustadas por la evolución.

La relación existente entre estabilidad y función ha sido objeto de estudio en diversos sistemas. En homeodominios, dicha conexión resulta especialmente intrigante si tenemos en cuenta su peculiar modo de interacción con el ADN.

Así, se ha demostrado que la formación del complejo lleva asociado el plegamiento de regiones parcialmente desordenadas del polipéptido (proceso conocido con el nombre de “ajuste inducido”), pudiendo involucrar en parte al propio núcleo estructurado del mismo. En casos extremos se ha llegado incluso a sugerir que la proteína únicamente adoptaría su estructura nativa como consecuencia del proceso de reconocimiento (Ippel *et al.*, 1998). Con independencia de su posible significado funcional, parece claro que el proceso de asociación requiere de un grado significativo de flexibilidad conformacional en el polipéptido libre, que puede implicar la desestructuración parcial o total de amplias regiones del mismo. En proteínas de reducido tamaño, dicha propiedad estaría normalmente asociada a una baja estabilidad.

En la presente memoria nos hemos propuesto, como objetivo principal, analizar la conexión existente entre estabilidad y función en homeodominios. Para ello, la aproximación experimental seleccionada se basa en el diseño, producción y análisis de diversos mutantes del sistema natural. En el camino, hemos tenido la oportunidad de abordar diversos objetivos secundarios, tales como evaluar la efectividad estabilizadora de los puentes salinos conservados en homeodominios en relación con las interacciones hidrófobas, entre otros. La relevancia de los resultados obtenidos se discute punto por punto a continuación.

4.1

ESTUDIO DE LA ESTABILIDAD DE LAS VARIANTES DEL HOMEODOMINIO DE HESX-1: PUENTES SALINOS FRENTE A PARES HIDRÓFOBOS

Nuestro análisis de secuencias de homeodominios ha puesto de manifiesto la existencia de tres puentes salinos interhélice conservados, que conectan los pares 19-30, 31-42 y 17-52. Dichas posiciones pueden albergar también pares hidrófobos, si bien, tal y como revela nuestro estudio, esta situación es, con mucho, menos frecuente en sistemas naturales. Esta observación sugiere que los citados puentes salinos podrían desempeñar papeles importantes, bien en la estabilización del polipéptido o bien en el reconocimiento de ADN. Las variantes de HESX-1 analizadas en el presente trabajo nos han dado la oportunidad de comparar el efecto estabilizador de uno u otro tipo de interacción. Los resultados obtenidos, detallados en la sección 3.1, nos permiten extraer las conclusiones que se abordan en los siguientes apartados.

4.1.1. Los puentes salinos en posiciones 19-30 y 31-42 resultan desestabilizantes respecto a pares hidrófobos, mientras que el efecto contrario se observa para el par 17-52

Las variantes de HESX-1 producidas fueron sometidas a un cuidadoso estudio de estabilidad. De él se derivaron los resultados detallados en los apartados 3.1.4, 3.1.6, 3.1.8, 3.1.10 y 3.1.11, y que se resumen a continuación:

a) Por lo que respecta a las posiciones fronterizas 19-30 (y adicionalmente 15), la sustitución de los residuos alifáticos presentes en el sistema natural por puentes salinos conlleva importantes descensos de estabilidad, con independencia de la polaridad de éstos últimos. El efecto estabilizante de las cadenas laterales alifáticas se manifiesta incluso al compararlo con aquellas variantes en las que, en principio, sería posible el establecimiento de una red cooperativa de puentes salinos conectando las posiciones 19-30-15-37.

b) En la misma línea, nuestros datos reflejan una clara preferencia por residuos alifáticos en las posiciones fronterizas 31 y 42. Sin embargo, en este caso la tendencia es mucho más acentuada. De hecho, la sustitución del puente salino R31-E42 por el par alifático L31-L42 en la variante **31**, nos lleva a una versión termófila del homeodominio de HESX-1, con temperaturas de desnaturalización en torno a los 90 °C.

c) En los casos a) y b), los datos obtenidos sugieren que las desestabilizaciones relativas promovidas por los residuos cargados son de origen fundamentalmente no electrostático ($\Delta G_{\text{no polar}}$ desfavorable).

d) En contraste con los resultados obtenidos para las posiciones 19-30 y 31-42, el par 17-52, caracterizado por un alto grado de exposición al disolvente, sólo parece tolerar residuos polares.

La [Figura 4.1](#) ilustra los puntos a) y b), mostrando la progresiva estabilización del polipéptido al reemplazar puentes salinos por pares alifáticos en las posiciones fronterizas 19, 30, 15, 31 y 42, lo que conduce finalmente a la variante termoestable **31**. Una explicación plausible de este efecto es que los residuos alifáticos V19, I30 y V15 (presentes en la proteína salvaje), así como L31 y L42 (variante **31**), proporcionan un empaquetamiento más optimizado con el núcleo hidrófobo que las cadenas laterales de E o R. Esto justificaría las diferencias encontradas en el término $\Delta G_{\text{no polar}}$. La reorganización local del núcleo hidrófobo observada en la estructura 3D obtenida por RMN para el mutante termófilo **31** (descrita en el apartado 3.1.7) es consistente con estos resultados. Alternativamente, el término $\Delta G_{\text{no polar}}$ podría reflejar en parte efectos electrostáticos poco sensibles a la sal, tales como la desolvatación incompleta de los residuos cargados en posiciones parcialmente expuestas al disolvente (Martin *et al.*, 2002).

Como excepción al comportamiento general descrito, un número reducido de variantes presentan contribuciones electrostáticas ($\Delta G_{\text{coulomb}}$) particularmente desfavorables. Casi todas ellas incorporan como rasgo común la mutación K37E. En nuestra opinión, la explicación más probable de este efecto es que dicha sustitución elimina un puente salino expuesto D33-K37, introduciendo en su lugar una interacción netamente repulsiva. Esta conclusión sería, a su vez, consistente con el análisis de los valores de pK_a correspondientes a los mutantes **12** y **15**, descritos en la presente memoria.

La supresión del puente salino 33-37 podría estar en el origen de la patología asociada a la mutación R37L en el homeodominio de SHOX. Tal y como se ha indicado en la introducción, dicha sustitución causa discondrosteosis (Grigelson *et al.*, 2000). Comparando las variantes **26** y **21** de nuestro trabajo, comprobamos que idéntica mutación en HESX-1 provoca una bajada de 7 °C en T_m .

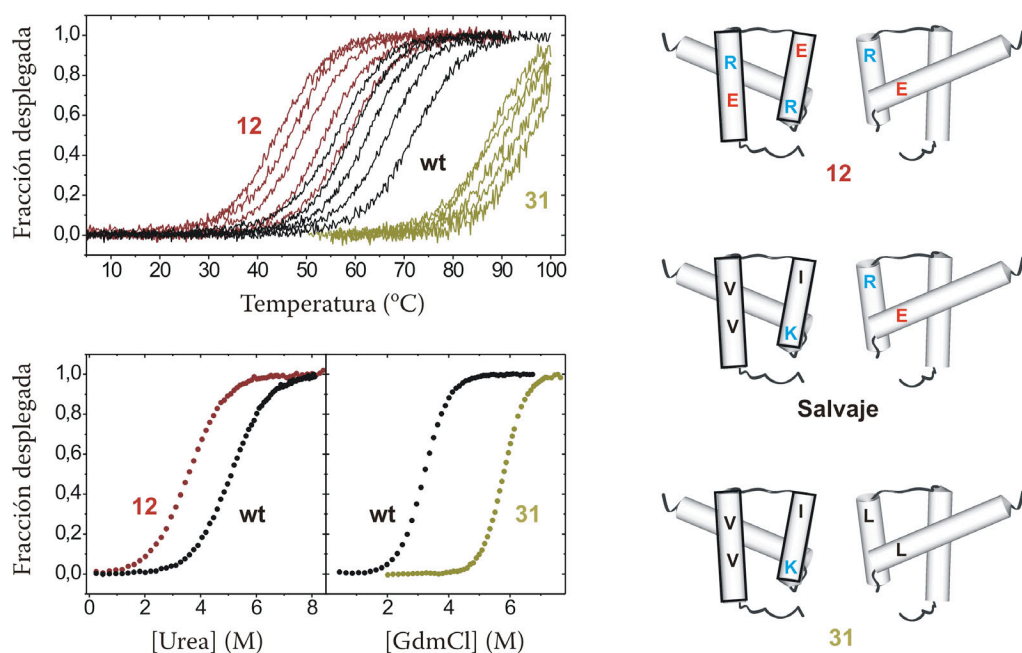


FIGURA 4.1: **Estabilización asociada a la reducción del número de puentes salinos** en posiciones 19-30, 15-37 y 31-42, ejemplificada con las variantes **12** (marrón), salvaje (negro) y **31** (dorado). Panel superior: Curvas de desnaturalización térmica registradas mediante dicroísmo circular, en tampón fosfato 10 mM pH 6,0, 0,5 mM DTT, y concentraciones crecientes de NaCl: de izquierda a derecha para cada variante, 0, 0,1, 0,4, 1 y 2 M NaCl. Panel inferior: Curvas de desnaturalización química con urea (izquierda) o GdmCl (derecha) registradas mediante dicroísmo circular, a 20 °C en tampón fosfato 10 mM pH 6,0, 0,5 mM DTT, 100 mM NaCl. A la derecha se muestra una representación esquemática de las tres proteínas. Los residuos cargados negativamente se muestran en rojo, y los cargados positivamente en azul.

En principio, la situación descrita para el par 33-37 se asemeja a la de las posiciones 17-52, también bastante expuestas al disolvente. La bajada de 15 °C en T_m medida para el mutante **33** (Y17-Y52), así como los problemas encontrados en su purificación y en la de la variante **32** (L17-Y52), son consistentes con una fuerte desestabilización asociada a la eliminación del puente salino E17-R52. Los bajos valores de pK_a^F determinados para E17 avalan la contribución favorable de esta carga negativa a la estabilidad del sistema. Por el contrario, residuos hidrófobos, tales como la tirosina o la leucina, en una localización tan expuesta, quedarían mal posicionados para empaquetar favorablemente con el núcleo de la proteína, y podrían afectar a la solubilidad del polipéptido, promoviendo su agregación. Estos problemas se verían acentuados por el comportamiento dinámico del polipéptido en la región en que se en-

cuentra dicho puente salino, el inicio de la hélice III-b, que tal y como se ha mencionado anteriormente está parcialmente desordenada.

Una revisión de la literatura pone de manifiesto el gran interés despertado por el puente salino 17-52 en otros sistemas. Así, un estudio teórico del homeodominio de Paired (Iurcu-Mustata *et al.*, 2001) muestra que la sustitución E17F produce una marcada desestabilización de 1,3 kcal/mol, muy en la línea de nuestros resultados. El efecto negativo observado al eliminar el puente salino 17-52 está también de acuerdo con evidencias experimentales obtenidas en otros homeodominios, que ya se han detallado previamente en la introducción. Así por ejemplo, destacan las estabilizaciones observadas en Engrailed (Stollar *et al.*, 2003) y en vnd (NK-2) (Weiler *et al.*, 1998; Gonzalez *et al.*, 2001) al introducir el puente salino mediante las mutaciones K52E y H52R, respectivamente. Es de destacar que esta última sustitución resulta tóxica en mosca (Fodor *et al.*, 2005). Por otro lado, cambios en esta interacción introducidos mediante mutaciones adecuadas en *even-skipped* (Frasch *et al.*, 1988), o *rough* (Heberlein *et al.*, 1994), provocan la acumulación de defectos en el desarrollo de *Drosophila* al aumentar la temperatura. Se ha descrito que la estabilización aportada por este puente salino suele venir acompañada de una mayor estructuración de la hélice III-b (Weiler *et al.*, 1998; Fausti *et al.*, 2001; Gonzalez *et al.*, 2001). Teniendo todos estos datos en cuenta, no resulta sorprendente comprobar que la posición 52 es uno de los puntos de mayor incidencia de mutaciones causantes de enfermedades humanas. En la introducción se han indicado varias que destacan por conllevar la rotura del puente salino 17-52 (Wu *et al.*, 1998; Swaroop *et al.*, 1999; Perveen *et al.*, 2000). De acuerdo con los argumentos propuestos en el presente trabajo, esto podría impedir al homeodominio alcanzar un plegamiento mínimamente estable, soluble y funcional. Los descensos de afinidad por ADN que se describen en algunos de los trabajos citados (Wu *et al.*, 1998; Swaroop *et al.*, 1999; Stollar *et al.*, 2003) serían consistentes con esta explicación.

Estas observaciones avalan la importancia del puente salino 17-52 en la estabilización del homeodominio, y explican su alto grado de conservación en secuencias naturales.

4.1.2. Discusión global sobre la aportación a la estabilidad del homeodominio de los puentes salinos respecto a pares hidrófobos. Influencia del grado de exposición al disolvente de los residuos involucrados

Diversos estudios han resaltado que la acumulación de puentes salinos, especialmente cuando se organizan en redes cooperativas, es una de las estrategias

más utilizadas en la adaptación de proteínas a altas temperaturas (Perutz y Raidt, 1975; Goldman, 1995; Karshikoff y Ladenstein, 2001; Pack y Yoo, 2003). Una de las razones que justifican esta observación experimental es que su fortaleza aumenta con la temperatura, mientras que lo contrario sucede con las interacciones puramente hidrófobas. Ello es debido a que la constante dieléctrica del agua disminuye al subir la temperatura, mientras que la del interior proteico se mantiene prácticamente igual (Karshikoff y Ladenstein, 2001; Olufsen *et al.*, 2007). En consecuencia, a altas temperaturas se potencia la interacción electrostática entre los residuos cargados a la vez que disminuye su coste de desolvatación. Por otro lado, la rotura de un puente salino requiere la superación de una energía de activación que nuevamente aumenta al elevar la temperatura, y que es mucho menor en el caso de las interacciones hidrófobas (Elcock, 1998).

Estas consideraciones contrastan con los resultados descritos en la presente memoria. De hecho, de acuerdo con el análisis que hemos llevado a cabo, de los tres puentes salinos conservados en homeodominios, dos (19-30 y 31-42) son claramente desestabilizantes con respecto a pares alifáticos como los que se pueden encontrar presentes en una pequeña fracción de secuencias. La [Figura 4.1](#) resume estos resultados, mostrando un acusado aumento en la estabilidad del homeodominio de HESX-1 al ir retirando puentes salinos de su estructura, lo que conduce a la variante termoestable **31**. Varias son las razones que podrían justificar esta tendencia general:

- En primer lugar, es muy importante considerar la geometría del puente salino. Factores como la distancia entre los grupos cargados, o el ángulo formado entre sus cadenas laterales, condicionan significativamente su contribución neta (Barril *et al.*, 1998; Kumar y Nussinov, 1999, 2001) y podrían no estar convenientemente optimizados en el homeodominio.
- En segundo lugar, la energía aportada por un puente salino depende, en buena medida, del ambiente electrostático generado por los residuos vecinos (Serrano *et al.*, 1990; Kumar y Nussinov, 1999; Xiao y Honig, 1999). Entornos desfavorables podrían contrarrestar la interacción atractiva entre las cadenas laterales cargadas, arrojando un efecto netamente desestabilizante para la interacción. Podría, por tanto, especularse con que el entorno de los pares 19-30 y 31-42 en HESX-1 proporciona interacciones repulsivas de calado con residuos cargados que se localizasen en estas posiciones. Sin embargo, un examen detallado de los datos experimentales resta fuerza a esta hipótesis. De hecho, la variación en la polaridad del puente salino 19-30, que implica un intercambio en el entorno de los residuos aniónico y catiónico, tiene un efecto relativamente menor sobre la energía libre de plegamiento. En otras palabras,

los puentes salinos en esta posición son en todos los casos desestabilizantes con respecto a pares alifáticos, independientemente de su polaridad. Por otro lado, una inspección de homeodominios naturales que albergan puentes salinos en estas posiciones (tales como Engrailed o Antennapedia), no avala la existencia de un entorno específicamente optimizado a tal efecto.

- En tercer lugar, es preciso considerar el grado de exposición al disolvente de cada uno de los puentes salinos analizados en la presente memoria. La formación de un puente salino durante el plegamiento de una proteína lleva asociado un significativo coste energético procedente de la desolvatación parcial de los residuos cargados. Diversos estudios teóricos muestran como, frecuentemente, dicho coste contrarresta la energía de atracción entre las cadenas laterales, arrojando un balance energético netamente desfavorable (Hendsch y Tidor, 1994; Barril *et al.*, 1998; Kumar y Nussinov, 1999). Este efecto cobra especial importancia para puentes salinos poco expuestos al disolvente, ya que entonces la desolvatación es más acusada. Una inspección detallada de los datos experimentales descritos en la presente memoria avala la relevancia de este fenómeno en el caso de HESX-1. De hecho, puede apreciarse como la sustitución de pares hidrófobos por puentes salinos tiene un efecto negativo sobre la estabilidad cuando involucra posiciones fronterizas (y por tanto parcialmente enterradas, como 19-30 ó 31-42) pero no cuando afecta a posiciones más expuestas y accesibles (tales como 17-52). Conviene recordar, tal y como se ha mencionado anteriormente, que el término $\Delta G_{\text{no polar}}$ empleado en nuestro análisis podría incluir parte del coste energético asociado a la desolvatación de las cadenas laterales (Martin *et al.*, 2002).

En resumen, pares alifáticos en las posiciones fronterizas 19-30 y 31-42, probablemente son capaces de establecer interacciones de empaquetamiento más optimizadas con el núcleo hidrófobo del homeodominio (lo cual se reflejaría en valores de $\Delta G_{\text{no polar}}$ más favorables). Adicionalmente, puentes salinos en estas posiciones parcialmente enterradas se verían penalizados por un elevado coste de desolvatación (parcialmente reflejado, nuevamente, en los términos $\Delta G_{\text{no polar}}$). Los argumentos opuestos podrían aplicarse al par 17-52. Estas consideraciones están de acuerdo con los resultados de otros grupos para diferentes sistemas (Waldburger *et al.*, 1995; Makhatadze *et al.*, 2003) y se corresponden bien con las predicciones teóricas de Hendsch y Tidor (Hendsch y Tidor, 1994).

Es preciso mencionar que, a pesar de reflejar el comportamiento de un único sistema (HESX-1), es probable que las conclusiones descritas en esta tesis sean generales para homeodominios, especialmente teniendo en cuenta que las propiedades estructurales de estas proteínas, así como su distribución de

cargas, están bastante conservadas. De hecho, las conclusiones son consistentes con nuestro análisis de 756 secuencias de homeodominios recogidas en el “Homeodomain Resource”, descrito en el apartado 3.1.2.

Finalmente, de acuerdo con lo anteriormente expuesto, cabe esperar que la influencia de los puentes salinos sobre la estabilidad de las proteínas sea altamente dependiente del contexto. El balance entre los factores considerados, que a su vez varían con la temperatura, el pH o la concentración de sales, es diferente según el caso estudiado. De hecho, es posible encontrar ejemplos de todo tipo de puentes salinos en lo referente a su aportación a la estabilidad. Así pues, las estimaciones de sus contribuciones energéticas van desde netamente estabilizantes (Marqusee y Baldwin, 1987; Lyu *et al.*, 1992; Marqusee y Sauer, 1994; Pervushin *et al.*, 1996; Spek *et al.*, 1998; Strop y Mayo, 2000) hasta insignificantes (Horovitz *et al.*, 1990; Serrano *et al.*, 1990; Sali *et al.*, 1991; Strop y Mayo, 2000) o incluso desestabilizantes (Hendsch y Tidor, 1994; Waldburger *et al.*, 1995; Marti y Bosshard, 2003).

4.1.3. Funciones de los puentes salinos que justifican su alta frecuencia de aparición en proteínas naturales

Los resultados recogidos en esta memoria, así como los descritos por otros autores, permiten plantear una pregunta de relevancia acerca de los puentes salinos: ¿por qué, pese a tener en muchos casos una influencia tan modesta o incluso negativa en la estabilidad de las proteínas, la naturaleza ha mantenido este tipo de interacciones? Existen varios motivos que podrían justificar esta selección:

- En primer lugar, debemos tener en cuenta que la maximización de la estabilidad no es, en ningún caso, un objetivo primordial para la naturaleza. En lugar de ello, dicho parámetro parece haber sido finamente ajustado por la evolución para cada sistema, en un valor óptimo para el desempeño de su función biológica. Existen dos razones fundamentales por las cuales una estabilidad excesiva podría tener efectos adversos sobre la función de una proteína. Por un lado, unas energías de plegamiento muy favorables alargarían excesivamente el tiempo de vida de la proteína en el medio biológico. Tal y como se ha comentado anteriormente, esto sería especialmente contraproducente en el caso de los factores de transcripción, para los cuales es deseable un rápido reciclaje. No resulta sorprendente, por tanto, comprobar que los homeodominios se caracterizan por valores de estabilidad marginales. Por otro lado, una excesiva estabilidad restaría flexibilidad a las proteínas, propiedad

que muchas veces es esencial para el correcto desempeño de sus funciones (Ó Fágáin, 1995; Freire, 2001; Teilum *et al.*, 2011). La incorporación de puentes salinos parcialmente enterrados podría ser utilizada por la naturaleza, en estos casos, como un recurso simple para limitar o reducir la energía libre asociada al plegamiento (Kajander *et al.*, 2000).

- En segundo lugar, la incorporación de puentes salinos puede ser importante para incrementar la especificidad del plegamiento. Es preciso considerar que, en general, existe un número limitado de conformaciones capaces de acomodar adecuadamente grupos cargados. Por el contrario, existen muchas formas energéticamente equivalentes de acomodar grupos hidrófobos (Lim y Sauer, 1991; Waldburger *et al.*, 1995; Wimley *et al.*, 1996). De acuerdo con esto, los puentes salinos podrían reducir el número de plegamientos estables compatibles con una secuencia dada. Adicionalmente, podrían dirigir el plegamiento, reduciendo la búsqueda de caminos posibles, aportando un beneficio cinético a costa de limitar la estabilidad termodinámica (Barril *et al.*, 1998).

- En tercer lugar, debe tenerse en cuenta que la presencia de cargas superficiales es primordial para incrementar la solubilidad de las proteínas en el medio acuoso y limitar así su agregación (Bosshard *et al.*, 2004). Esto sería especialmente crítico en el interior celular donde existe una alta concentración de proteínas. La incorporación de puentes salinos parcial o totalmente expuestos permitiría satisfacer este requerimiento sin incrementar la carga neta del sistema.

- Finalmente, existen casos en que los puentes salinos aportan grupos cargados relevantes para la función de la proteína. A modo de ejemplo, estas interacciones podrían proporcionar un entorno electrostático adecuado para la catálisis enzimática, o posicionar cargas para la interacción con un ligando (Bosshard *et al.*, 2004). En este sentido, hemos destacado la importancia de los residuos R31 y R52 del homeodominio de HESX-1, componentes de sendos puentes salinos, en la interacción con ADN.

Como conclusión, podemos afirmar que los puentes salinos son herramientas utilizadas por la naturaleza para lograr un amplio abanico de propósitos, que no siempre tienen que ver con aumentar la estabilidad de la proteína. Para este cometido, nuestro análisis muestra que, en muchos casos, las interacciones hidrófobas y de van der Waals son más apropiadas.

4.1.4. Correlación entre la posición 26 y la naturaleza del par 19-30: influencia de la tríada 19-26-30 en la estabilidad del homeodominio

Nuestro estudio de estabilidad, en combinación con el análisis comparativo del “Homeodomain Resource” llevado a cabo, revela los resultados detallados en el apartado 3.1.5. Las conclusiones obtenidas se resumen en los siguientes puntos:

a) En secuencias naturales, la presencia de prolina en posición 26 (perteneiente al lazo I-II) puede venir acompañada de residuos tanto alifáticos como polares en 19-30. Esta observación está de acuerdo con la viabilidad de las variantes salvaje (V19/P26/I30) y **19** (E19/P26/R30) descritas en la presente memoria.

b) Por otro lado, la variante **17** (E19/L26/R30), pese a presentar la combinación de residuos más frecuente en homeodominios, es, curiosamente, ligeramente menos estable que la variante **19** (E19/P26/R30).

c) Por último, una cadena lateral de leucina en posición 26, más grande y móvil que la prolina, resulta incompatible con cadenas laterales alifáticas en las posiciones 19-30. De hecho, dicha combinación de residuos no está presente en ninguna secuencia natural. La mutación P26L introducida en el polipéptido salvaje, para dar la variante **1** (V19/L26/I30), nos permite comprobar cómo violaciones de esta regla conducen a una severa desestabilización o agregación del homeodominio.

La conclusión b) estaría de acuerdo con los resultados descritos para otras proteínas, de acuerdo con los cuales sustituciones de prolina por otro residuo acostumbran a causar claras bajadas de estabilidad (Mayr *et al.*, 1993). Parece lógico suponer que la influencia favorable de P26 sobre la estabilidad conformacional del polipéptido está, en mayor o menor medida, relacionada con el mayor grado de preorganización y menor movilidad interna de este residuo. De acuerdo con esta idea, el carácter cíclico de su cadena lateral tendría como consecuencia una menor pérdida de grados de libertad conformacional asociada al plegamiento. Además, en consonancia con nuestros resultados, se han observado casos de correlación entre termoestabilidad proteica y abundancia de prolina en lazos (Delboni *et al.*, 1995; Ó Fágáin, 1995).

Teniendo en cuenta la mayor frecuencia de aparición en la naturaleza de la combinación menos estable, E19/L26/R30, podría especularse con un cierto papel funcional para la misma. De hecho, se ha propuesto que esta tríada participa en una red de interacciones que involucra residuos claves en el reconocimiento de ADN (Clarke, 1995; Stollar *et al.*, 2003). No obstante, los estudios de asociación descritos en la presente memoria (ver final del apartado 3.2.2), parecen desmentir este supuesto.

El comportamiento observado para la variante **1** (V19/L26/I30), señalado en la conclusión c), podría resultar sorprendente si consideramos que P26L es una mutación conservativa a simple vista (especialmente teniendo en cuenta que muchos homeodominios tienen leucina en posición 26). Sin embargo, estaría de acuerdo con estudios previos de diseño de proteínas. Así, se ha demostrado que, si bien se permiten algunas variaciones en posiciones enterradas de una proteína, existen límites en las secuencias que resultan en un plegamiento estable y similar al nativo (Lim y Sauer, 1991; Lim *et al.*, 1992; Dahiyat y Mayo, 1997). La omisión de una mínima especificidad en las interacciones de empaquetamiento puede introducir desorden y una movilidad conformacional excesiva en proteínas de diseño (Dahiyat y Mayo, 1997). Teniendo esto en cuenta, un empaquetamiento específico entre los residuos 19, 26 y 30 parece crítico para la estabilidad del homeodominio.

La importancia de la posición 26 y su red de interacciones con residuos vecinos ha sido puesta de relieve por el efecto patológico de la mutación P26L en el homeodominio de ARX humano, anteriormente descrito en la introducción (Scheffer *et al.*, 2002; Strømme *et al.*, 2002). Merece la pena considerar que dicha sustitución no afecta a posiciones directamente involucradas en la unión a ADN, a diferencia de la mayoría de las mutaciones causantes de enfermedades en homeodominios. Se ha sugerido, no obstante, que su presencia en ARX puede ser inapropiada en el contexto de los residuos vecinos 19 y 30, afectando a la propensión del homeodominio para el ajuste inducido (Stollar *et al.*, 2003). Sin embargo, nuestros resultados, y en particular los presentados para el mutante **1** (V19/L26/I30), sugieren una explicación alternativa: que la sustitución P26L desestabiliza gravemente la conformación nativa. Otra enfermedad humana destacada en la introducción es aquella causada por la sustitución P26S en POU1F1 (Pernasetti *et al.*, 1998). Nuestros resultados sugieren que la explicación a la disfunción subyacente también podría encontrarse en la sensibilidad que la estabilidad del homeodominio exhibe ante los cambios en el residuo 26.

4.1.5. Los metilenos de las cadenas laterales involucradas en puentes salinos contribuyen significativamente a la estabilidad del plegamiento

Los residuos cargados en posiciones 19, 30, 31 y 42 pueden contribuir positivamente a la estabilidad del homeodominio no sólo a través de interacciones electrostáticas, sino mediante el empaquetamiento de sus grupos metileno con el núcleo hidrófobo de la proteína (lo que incluye interacciones de van der Waals más el efecto hidrófobo). Estos contactos podrían ser especialmente

importantes para los pares 19-30 y 31-42, debido a su particular localización fronteriza en la estructura del homeodominio. Los resultados obtenidos, detallados en el apartado 3.1.9, confirman esta hipótesis de partida. Así, las desestabilizaciones inducidas por sustituciones simples E→D (que representan la eliminación de un único grupo metileno) en los mutantes de HESX-1 **2** (D19/R30) y **27** (R31/D42) con respecto a las variantes **19** (E19/R30) y salvaje (R31/E42), resultaron ser de 1,7 y 0,7 kcal/mol, respectivamente. Además, de acuerdo con nuestra aproximación, la práctica totalidad de dichos valores fue debida a contribuciones apolares.

Estas observaciones se mostraron consistentes con los datos presentes en la literatura. A modo de ejemplo, la sustitución E→D en un puente salino lisina-glutámico de la ubiquitina, también establecido entre posiciones fronterizas (Makhatadze *et al.*, 2003), conlleva una desestabilización de 1,0 kcal/mol. Sin embargo, en este caso los autores justifican el resultado en base a la menor atracción electrostática entre lisina y ácido aspártico, motivada por el alejamiento entre cargas que supone acortar la cadena lateral del residuo ácido. Existen a su vez diversos ejemplos en la literatura de sustituciones I→V consistentes, al igual que en nuestro caso, en la eliminación de un grupo metileno (Shortle *et al.*, 1990; Serrano *et al.*, 1992; Jackson *et al.*, 1993; Takano *et al.*, 1995). Dichas mutaciones conllevan desestabilizaciones de entre 0,5 y 1,75 kcal/mol. La contribución al efecto hidrófobo aportada por el metileno extra de E respecto a D queda ilustrada también en la energía requerida para la transferencia de estos aminoácidos de ciclohexano a agua, que es 1,9 kcal/mol mayor para E (Fersht, 1999). En general, todos estos valores guardan buena correspondencia con los resultados descritos en la presente memoria, y nos dan una idea de lo que un solo grupo metileno puede aportar a la estabilidad de una proteína.

4.2

ESTUDIO DE LAS PROPIEDADES DE UNIÓN A ADN DE UNA SELECCIÓN DE VARIANTES DE HESX-1: RELACIÓN ENTRE ESTABILIDAD Y FUNCIÓN EN HOMEODOMINIOS Y SU IMPORTANCIA EVOLUTIVA

Los trabajos descritos en la presente memoria han conducido a la preparación de homeodominios con un amplio rango de estabilidades. Es de destacar que, en muchos de los casos, dichas variantes se han obtenido mediante la incorporación de sustituciones presentes en otros homeodominios naturales y localizadas en posiciones lejanas a la superficie de interacción proteína-ADN (tales como 19-30 y 15-37). De acuerdo con esto, cabría esperar que no tuvieran una incidencia directa sobre la asociación. A pesar de ello, podrían repercutir sobre la estabilidad del complejo de forma indirecta, modulando la flexibilidad conformacional del polipéptido y afectando al proceso de ajuste inducido. En principio, cabría esperar que dicho efecto fuese especialmente acentuado si, tal y como se ha propuesto, el ajuste inducido involucrase no sólo las regiones desordenadas sino también en buena medida el núcleo estructurado del homeodominio. Con objeto de comprobar esta posibilidad, como último paso de nuestro trabajo se analizó la asociación específica de una selección de mutantes a ADN empleando microcalorimetría (ITC).

Como objetivo secundario, este estudio nos permitió, adicionalmente, analizar el papel de puentes salinos conservados, tales como 31-42, en el reconocimiento de ADN. A continuación se discuten los resultados obtenidos, empezando por este último punto.

4.2.1. El puente salino 31-42, altamente conservado en homeodominios, tiene un papel modesto en el reconocimiento de ADN

El papel de las posiciones 31 y 42 en la unión a ADN ha sido destacado por numerosos estudios estructurales (Kissinger *et al.*, 1990; Billeter *et al.*, 1993; Grant *et al.*, 2000). Su relevancia funcional se ve, a su vez, apoyada por el hecho de que mutaciones en ambas posiciones han sido asociadas a diversas enfermedades de carácter hereditario (Vastardis *et al.*, 1996; Freund *et al.*, 1997; Kulak

et al., 1998). Pese a ello, el trabajo descrito en la presente memoria constituye el primer análisis cuantitativo del papel desempeñado por ambos residuos en la función del homeodominio. Así, las valoraciones por ITC llevadas a cabo con las variantes salvaje (R31/E42), **28** (R31/L42), **29** (L31/E42) y **31** (L31/L42) nos han permitido conocer el efecto de R31 y de su puente salino con E42 en el reconocimiento de ADN. Los resultados obtenidos, detallados en el apartado 3.2.1, se resumen en dos:

a) La posición 31 del homeodominio está involucrada en la estabilización del complejo proteína/ADN. Sin embargo, su papel es relativamente modesto (un factor de 4,4 veces en K_b).

b) La posición 42 no parece jugar ningún papel en la asociación, ni directamente (mediante interacciones electrostáticas repulsivas con el ADN) ni a través de su interacción con R31 (reduciendo la movilidad interna de este residuo o preorganizándolo para su interacción con el ADN).

La relativamente moderada influencia de R31 en el proceso de asociación resulta ciertamente llamativa. De hecho, se han descrito efectos muy superiores asociados a la supresión de interacciones arginina-fosfato en otros sistemas. A modo de ejemplo, en homeodominios, la sustitución R53C en HESX-1 causa una pérdida de afinidad de más de 1000 veces (de la Mata *et al.*, 2002), y la mutación R53H en Proboscipedia conlleva una caída de 20 veces en K_b (Benassayag *et al.*, 1997). En otros factores de transcripción, encontramos casos de mutaciones equivalentes en Max p21, con bajadas de afinidad de entre 15 y 200 veces (Meier-Andrejszki *et al.*, 2007), o en p53, con caídas de más de 100 veces (Bullock *et al.*, 2000). Estos ejemplos muestran la importancia de las interacciones arginina-fosfato en la unión de proteínas a ADN. Sin embargo, el contraste con nuestros resultados subraya hasta qué punto el efecto estabilizante de las mismas es dependiente de contexto.

En base a estos resultados, puede resultar sorprendente la abundancia de enfermedades humanas heredables asociadas a mutaciones en los residuos R31 o E42 en homeodominios como los de MSX-1, PITX2 (RIEG1) o CRX (Vastardis *et al.*, 1996; Freund *et al.*, 1997; Kulak *et al.*, 1998). De hecho, como comentábamos en la introducción, la posición 31 es una de las localizaciones de mayor incidencia de mutaciones patológicas. Es posible, por tanto, que a pesar de tener un efecto relativamente modesto sobre las propiedades de unión, el residuo R31 sea, de hecho, importante en la consecución de un evento regulador preciso. Sin embargo, a la luz de los resultados descritos en la presente memoria, también podría plantearse que las sustituciones en 31 y 42 interfiriesen con la función del homeodominio afectando de modo preferente a su estabilidad (y en último término a la cantidad de proteína funcional). De hecho,

ambas posiciones parecen tener una influencia significativamente mayor sobre este parámetro que sobre el proceso de asociación. Una posibilidad interesante es que el enterramiento parcial de R31 tenga como misión limitar la estabilidad del homeodominio, manteniéndola en unos márgenes aceptables para el desempeño de su función. Esta hipótesis se vería reforzada por la observación de que, en principio, interacciones arginina-fosfato análogas a la establecida desde la posición 31 podrían darse desde posiciones próximas más expuestas al disolvente, como 28 ó 42. Presumiblemente, dichos contactos contribuirían a la asociación sin promover desestabilizaciones pronunciadas del homeodominio. Interacciones arginina-fosfato mediadas por las posiciones 28 ó 42 se observan en las estructuras cristalográficas de los complejos con ADN correspondientes a los homeodominios de HOXA9 y de MAT α 2 (Figura 4.2).

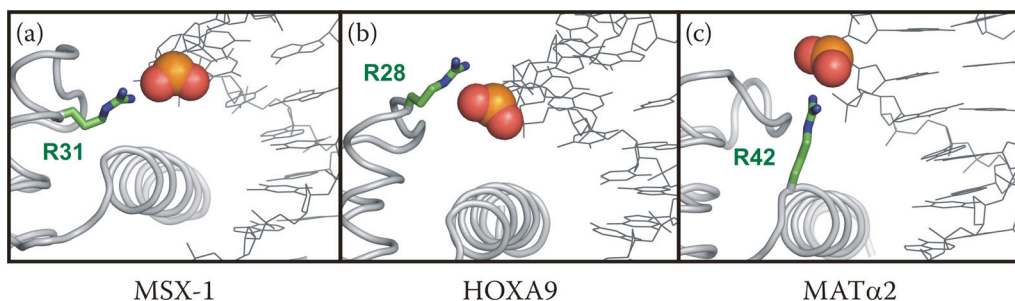


FIGURA 4.2: El contacto conservado R31-fosfato (representado en (a) para el homeodominio de MSX-1, código PDB: 1IG7) puede ser reemplazado en algunos casos por interacciones equivalentes como R28-fosfato (representado en (b) para el homeodominio de HOXA9, código PDB: 1PUF) o R42-fosfato (representado en (c) para el homeodominio de MAT α 2, código PDB: 1K61). En los tres paneles se representa la arginina en verde y se destaca el fosfato en esferas.

4.2.2. Las propiedades de unión a ADN de las variantes con sustituciones en las posiciones 19, 30, 15 y 37 revelan una independencia entre la estabilidad y la función del homeodominio

El análisis mediante microcalorimetría de las propiedades de unión a ADN de las variantes de HESX-1 con mutaciones en las posiciones 19-30 y 15-37 se detalla en el apartado 3.2.2. Merece la pena destacar que dicho análisis constituye el primer estudio detallado de homeodominios en el que se ensaya la influencia en la asociación de posiciones no involucradas directamente en contactos con el ADN. Las conclusiones obtenidas se resumen a continuación:

a) La influencia de la estabilidad del homeodominio en sus propiedades de unión a ADN es prácticamente despreciable. De hecho, las variantes analizadas presentan energías libres de plegamiento de entre $-2,01$ y $-4,60$ kcal/mol a $20\text{ }^{\circ}\text{C}$ (ver [Tabla 3.1](#)), mientras que todas ellas exhiben afinidades similares por ADN (los máximos valores de $\Delta\Delta G_b$ medidos no superan las $0,3$ kcal/mol). Por lo tanto, la estabilidad y la función no muestran correlación en el homeodominio de HESX-1. Estos datos parecen sugerir que la flexibilidad conformacional de las regiones estructuradas tiene una influencia menor o despreciable sobre las propiedades de unión del homeodominio.

b) El análisis de los valores de ΔH_b y $T\Delta S_b$ nos ha permitido sugerir que esta ausencia de correlación está motivada por un fenómeno de compensación entalpía-entropía (Dunitz, 1995; Cooper *et al.*, 2001).

c) La escasa variabilidad en los valores de $\Delta C_{p,b}$ que hemos medido para un conjunto reducido de variantes (salvaje, **16**, **11** y **7**), parece indicar que la magnitud del ajuste inducido es independiente de la estabilidad del homeodominio. Esta observación sugiere que el núcleo del homeodominio no juega un papel significativo en el mecanismo del ajuste inducido, el cual quedaría fundamentalmente restringido a las regiones desestructuradas del polipéptido (hélice III y brazo N-terminal).

Podemos encontrar algunos datos en consonancia con los nuestros en los tempranos estudios del grupo de Abate, en los cuales ciertas mutaciones en las hélices I y II no ejercían un efecto pronunciado en la afinidad por ADN del homeodominio de MSX-1, de acuerdo con experimentos de retardo en gel (EMSA) (Shang *et al.*, 1994; Isaac *et al.*, 1995). Precisamente, dichas mutaciones incluían la sustitución del puente salino K19-E30 por A19-A30 (Shang *et al.*, 1994), así como su cambio de polaridad para dar el par E19-R30 (Isaac *et al.*, 1995). Estos datos y los nuestros apuntan, por tanto, a una escasa influencia de este puente salino, y en general de las posiciones 19 y 30, en la unión del homeodominio a ADN. En consecuencia, el papel de estos residuos y su red de interacciones en el ajuste inducido, propuesto por el grupo de Fersht (Stollar *et al.*, 2003), no encuentra de momento apoyo experimental.

Nuestras conclusiones son probablemente significativas en el contexto de la evolución. De este modo, si los factores dinámicos del núcleo hidrófobo del homeodominio fuesen importantes para la formación del complejo, muchas de las mutaciones que afectasen a la estabilidad conformacional del homeodominio tendrían también un efecto en la unión a ADN. Por el contrario, la ausencia de correlación entre estabilidad y función proporcionaría una gran versatilidad adaptativa al permitir una *modulación independiente* de ambas propiedades. En este contexto, merece la pena considerar que los homeodominios están presentes en toda la escala de eucariotas, lo que abarca desde

peces que viven en aguas polares, hasta aves cuya temperatura corporal supera los 40 °C. La adaptación de la estabilidad a la temperatura del entorno es una condición ineludible en todas las proteínas, si bien es, como ya se ha señalado anteriormente, especialmente crítica en homeodominios. Además, esta minuciosa adaptación a entornos diversos no debe impedir a estas proteínas mantener una alta afinidad por ADN y especificidades bien definidas. Lo sorprendente es que los homeodominios cumplen con todos estos requisitos pese a exhibir una tremenda conservación de la secuencia que se traduce en un plegamiento invariable. Aunque meramente especulativo, la independencia relativa entre estabilidad y propiedades de asociación, puesta de manifiesto en el presente trabajo, podría facilitar en buena medida esta tarea.

En base al presente trabajo, podemos sugerir que los cambios de residuos en las posiciones 19-30 y 15-37 contribuyen a ajustar la estabilidad térmica del homeodominio para unas condiciones particulares, manteniendo inalteradas sus propiedades de asociación. Puesto que, según nuestros resultados, el par 31-42 tampoco tiene una influencia crítica en el reconocimiento de ADN, estos residuos podrían también jugar un papel modulador parecido al de 19-30 y 15-37. En apoyo de esta propuesta, el análisis de la información de secuencias de homeodominios que hemos llevado a cabo, revela que únicamente en animales de sangre caliente es posible encontrar pares hidrófobos en posiciones 31-42, altamente estabilizantes según nuestros resultados. Además, las secuencias de animales de sangre caliente exhiben también una abundancia de puentes hidrófobos entre las posiciones 19 y 30 significativamente mayor que las de animales de sangre fría, plantas y hongos (3,5% frente a 1%).

Como resumen global, la influencia de las posiciones 19, 30, 15, 37, 31 y 42 en las propiedades de unión a ADN del homeodominio van de moderadas (en el caso del residuo 31) a prácticamente despreciables. Por el contrario, hemos visto que estas posiciones tienen un efecto muy marcado en la estabilidad conformacional de la proteína. Teniendo esto en cuenta, proponemos que el papel principal de los puentes salinos o hidrófobos en estas posiciones es el ajuste de la estabilidad del homeodominio a un ambiente determinado. Al no ejercer una influencia significativa en la unión a ADN, su valor evolutivo reside, por tanto, en permitir una modulación independiente de la estabilidad de la proteína y de su energía de unión.

5 *PROCEDIMIENTOS EXPERIMENTALES*

5.1

OBTENCIÓN DE LAS VARIANTES DEL HOMEODOMINIO DE HESX-1

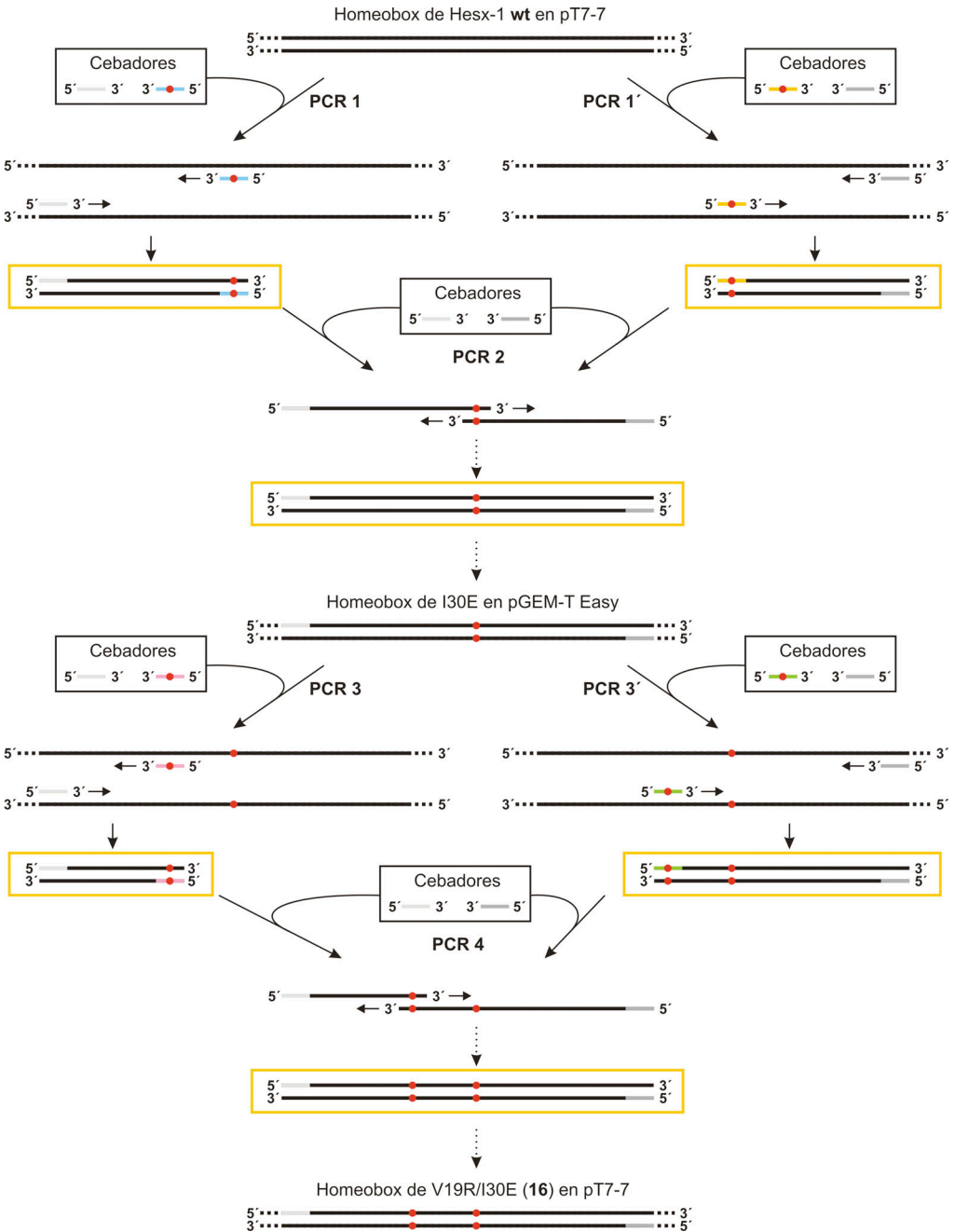
5.1.1. Construcción de plásmidos y mutagénesis dirigida

El cDNA correspondiente al homeobox salvaje de *Hesx-1* clonado en el vector de expresión pT7-7 (pT7-7-**WT**) fue cedido por el Dr. José Luis García (Centro de Investigaciones Biológicas, CSIC). El homeobox de la variante **16** (doble mutante V19R/I30E), fue obtenido a partir del plásmido pT7-7-**WT**, empleando el método de Higuchi *et al.* (Higuchi *et al.*, 1988), representado en la [Figura 5.1](#). Los cebadores usados como portadores de las mutaciones son los especificados en la [Tabla 5.1](#). Por otro lado, los cebadores utilizados para los extremos del homeobox fueron:

1: 5'-GGAATTCCATATGGGCCGAAGACCAAGAACTGC-3'

2: 5'-CGGGATCCTTACATTAGAAACTGTGATTCTC-3'

Todos ellos fueron sintetizados por el Servicio de Química de Proteínas del Centro de Investigaciones Biológicas (CSIC). Las dos mutaciones fueron introducidas secuencialmente. En primer lugar se introdujo la mutación I30E. En cada una de las dos primeras PCRs del método encaminadas a obtener esta mutación, se utilizaron 2 ng de plásmido pT7-7-**WT**, 1 nmol de cada cebador, 1mM dNTPs (MBI Fermentas) y 3mM MgCl₂, en un volumen de 100 µl. La ADN polimerasa empleada fue la Ecozyme (Ecogen) (2,5 U), que se caracteriza por una alta fidelidad de copia. Para la segunda PCR se utilizaron 2 ng de cada fragmento de ADN generado en las primeras PCRs, 25 pmol de cada cebador, 1mM dNTPs y 3mM MgCl₂, en un volumen de 35 µl. En este caso empleamos la EcoTaq (Ecogen) (1,75 U), enzima que añade un nucleótido de adenina en el extremo 3' de cada monohebra de ADN. Dicha extensión es necesaria para el clonaje subsiguiente que se explica a continuación. El fragmento de ADN obtenido (homeobox con la mutación I30E) fue ligado con el vector pGEM-T Easy (Promega), utilizando 50 ng de vector, 500 ng de inserto y 1,5 U



de T4 DNA ligasa incluida con el kit del pGEM-T Easy, en un volumen de 5 μ l. Con objeto de obtener cantidades manejables del plásmido pGEM-T Easy con el homeobox I30E, la mezcla de ligación fue utilizada para transformar células

FIGURA 5.1 (PÁGINA ANTERIOR): Esquema del método de Higuchi *et al.* de mutagénesis dirigida (Higuchi *et al.*, 1988), que ilustra la introducción de la doble mutación V19R/I30E en la secuencia del homeobox salvaje, para obtener la variante **16**. Los cebadores utilizados se representan con el siguiente código de colores (la numeración es la usada en el texto y en la [Tabla 5.1](#)): gris claro: extremo 1; gris oscuro: extremo 2; amarillo: I30E 1; azul: I30E 2; verde: V19R 1; rosa: V19R 2. La posición de las mutaciones se representa con un punto rojo. En aquellos casos en que el homeobox se encuentra dentro de un plásmido, éste se representa con pequeñas líneas discontinuas. Encuadrados en amarillo se muestran los productos finales de las distintas PCRs. Las flechas discontinuas indican etapas del proceso de PCR o bien del proceso de clonaje omitidas para dar claridad.

TABLA 5.1: Resumen de las reacciones de mutagénesis dirigida^a

Mutante	Plásmido de partida	Mutación	Cebador	Secuencia 5'→3'	Plásmido resultante
1	pT7-7- WT	P26L	1	GAGTAAACTGCTATCTGGGTATCGAT	pT7-7- 1
			2	ATCGATACCCAGATAGCAGTTTACTC	
2	pT7-7- 19	E19D	1	GAAGTGTAGAAAATGATTTTAGAGTAACTGC	pT7-7- 2
			2	GCAGTTTACTCTAAAATCATTTTCTAACACTTC	
3	pT7-7- 4	V15E	1	CAAAACCAGATTGAAGAAATTAGAAAATCGC	pT7-7-pre 3
	pT7-7-pre 3	K37R	2	GCGATTTTCTAATTCTTCAATCTGGTTTTG	pT7-7- 3
			1	GAAGACTTAGCTCAACGCTTGAATCTAGAGG	
			2	CCTCTAGATTCAAGCGTTGAGCTAAGTCTTC	
4	pT7-7- WT	V19R	1	GAAGTGTAGAAAATCGCTTTAGAGTAAAC	pT7-7- 4
			2	CAGTTTACTCTAAAGCGAATTTTCTAACAC	
5	pT7-7- 8	V15R	1	CAAAACCAGATTGAACGCTTAGAAAATGTC	pT7-7-pre 5
	pT7-7-pre 5	K37E	2	GACATTTTCTAAGCGTTCAATCTGGTTTTG	pT7-7- 5
			1	GAAGACTTAGCTCAAGAAATTGAATCTAGAGG	
			2	CCTCTAGATTCAATTCTTGAGCTAAGTCTTC	
6	pT7-7- 14	V15R	1	CTCAAAACCAGATTGAACGCTTAGAAAATGAATTTAG	pT7-7-pre 6
	pT7-7-pre 6	K37E	2	CTAAATTCATTTTCTAAGCGTTCAATCTGGTTTGAG	pT7-7- 6
			1	GAAGACTTCGCTCAAGAAATTGAATCTAGAGG	
			2	CCTCTAGATTCAATTCTTGAGCTAAGTCTTC	
7	pT7-7- 19	V15R	1	CTCAAAACCAGATTGAACGCTTAGAAAATGAATTTAG	pT7-7- 7
			2	CTAAATTCATTTTCTAAGCGTTCAATCTGGTTTGAG	
8	pT7-7- WT	I30R	1	GCTATCCTGGTATCGATCTAGAGAAGACTTCGC	pT7-7- 8
			2	GCTAAGTCTTCTCTACGATCGATACCACGATAGC	
9	pT7-7- 19	K37E	1	GAAGACTTAGCTCAAGAAATTGAATCTAGAGG	pT7-7- 9
			2	CCTCTAGATTCAATTCTTGAGCTAAGTCTTC	
10	pT7-7- WT	K37E	1	GAAGACTTAGCTCAAGAAATTGAATCTAGAGG	pT7-7- 10
			2	CCTCTAGATTCAATTCTTGAGCTAAGTCTTC	
11	pT7-7- 16	V15E	1	CAAAACCAGATTGAAGAAATTAGAAAATCGC	pT7-7- 11
			2	GCGATTTTCTAATTCTTCAATCTGGTTTTG	
12	pT7-7- 11	K37R	1	GAAGACTTAGCTCAACGCTTGAATCTAGAGG	pT7-7- 12
			2	CCTCTAGATTCAAGCGTTGAGCTAAGTCTTC	
13	pT7-7- 16	K37R	1	GAAGACTTAGCTCAACGCTTGAATCTAGAGG	pT7-7- 13
			2	CCTCTAGATTCAAGCGTTGAGCTAAGTCTTC	

CONTINÚA EN LA PÁGINA SIGUIENTE

TABLA 5.1 (CONT.)^a

Mutante	Plásmido de partida	Mutación	Cebador	Secuencia 5'→3'	Plásmido resultante
14	pT7-7-WT	V19E	1 2	GAAGTGTTAGAAAAATGAATTTAGAGTAAACTGC GCAGTTTACTCTAAATTCATTTTCTAACACTTC	pT7-7-14
15	pT7-7-7	K37E	1 2	GAAGACTTCGCTCAAGAAATTGAATCTAGAGG CCTCTAGATTCAATTCCTTGAGCTAAGTCTTC	pT7-7-15
16	pT7-7-WT	V19R	1 2	GAAGTGTTAGAAAAATCGCTTTAGAGTAAAC CAGTTTACTCTAAAGCGATTTTCTAACAC	pT7-7-16
		I30E	1 2	CTATCCTGGTATCGATGAAAGAGAAGACTTAG GAGCTAAGTCTTCTCTTTCATCGATACCAGG	
17	pT7-7-19	P26L	1 2	GAGTAACTGCATCTGGGTATCGAT ATCGATACCCAGATAGCAGTTTACTC	pT7-7-17
18	pT7-7-WT	V15R	1 2	CAAAACCAAGATTGAACGCTTAGAAAAATGTC GACATTTTCTAAGCGTTCAATCTGGTTTTG	pT7-7-18
19	pT7-7-8	V19E	1 2	GAAGTGTTAGAAAAATGAATTTAGAGTAAACTGC GCAGTTTACTCTAAATTCATTTTCTAACACTTC	pT7-7-19
20	pT7-7-18	K37E	1 2	GAAGACTTAGCTCAAGAAATTGAATCTAGAGG CCTCTAGATTCAATTCCTTGAGCTAAGTCTTC	pT7-7-20
21	pT7-7-WT	K37L	1 2	GAAGACTTAGCTCAATTATTGAATCTAGAG CTCTAGATTCAATAATTGAGCTAAGTCTTC	pT7-7-21
22	pT7-7-25	V15E	1 2	CAAAACCAAGATTGAAGAATTAGAAAAATGTC GACATTTTCTAATTCCTCAATCTGGTTTTG	pT7-7-pre22
	pT7-7-pre22	K37R	1 2	GAAGACTTAGCTCAACGCTTAGAATCTAGAGG CCTCTAGATTCAAGCGTTGAGCTAAGTCTTC	pT7-7-22
23	pT7-7-WT	V15E	1 2	CAAAACCAAGATTGAAGAAATTAGAAAAATGTC GACATTTTCTAATTCCTCAATCTGGTTTTG	pT7-7-23
24	pT7-7-26	V15E	1 2	CAAAACCAAGATTGAAGAATTAGAAAAATGTC GACATTTTCTAATTCCTCAATCTGGTTTTG	pT7-7-24
25	pT7-7-WT	I30E	1 2	CTATCCTGGTATCGATGAAAGAGAAGACTTAG GAGCTAAGTCTTCTCTTTCATCGATACCAGG	pT7-7-25
26	pT7-7-WT	K37R	1 2	GAAGACTTAGCTCAACGCTTAGAATCTAGAGG CCTCTAGATTCAAGCGTTGAGCTAAGTCTTC	pT7-7-26
27	pT7-7-WT	E42D	1 2	GCTCAAAAATTGAATCTAGAGGATGACAGAATCCAG CTGGATTCTGTCACTCCTAGATTCAATTTTGAGC	pT7-7-27
28	pT7-7-WT	E42L	1 2	GCTCAAAAATTGAATCTAGAGTTAGACAGAATCCAG CTGGATTCTGTCTAACTCTAGATTCAATTTTGAGC	pT7-7-28
29	pT7-7-WT	R31L	1 2	GGTATCGATATTTTGAAGACTTAGCTC GAGCTAAGTCTTCTTAAATATCGATACC	pT7-7-29
30	pT7-7-31	L31I	1 2	GGTATCGATATTAATTGAAGACTTAGCTC GAGCTAAGTCTTCAATAATATCGATACC	pT7-7-30
31	pT7-7-29	E42L	1 2	GCTCAAAAATTGAATCTAGAGTTAGACAGAATCCAG CTGGATTCTGTCTAACTCTAGATTCAATTTTGAGC	pT7-7-31
32	pT7-7-33	Y17L	1 2	CCAGATTGAAGTGTTATTAAATGTCTTTAGAG CTCTAAAGACATTTAATAACACTTCAATCTGG	pT7-7-32
33	pT7-7-WT	E17Y	1 2	CCAGATTGAAGTGTTATATAATGTCTTTAGAG CTCTAAAGACATTATATAACACTTCAATCTGG	pT7-7-E17Y
	pT7-7-E17Y	R52Y	1 2	GGTTTCAAAATTATCGTGCAAAACTG CAGTTTTGCACGATAATTTTGAAACC	pT7-7-33

^a El triplete portador de la mutación en cada cebador se muestra subrayado.

competentes (por CaCl_2) de la cepa DH5 α de *Escherichia coli* (Invitrogen). La correcta introducción de la mutación fue verificada por secuenciación en el Servicio de Secuenciación Automática de DNA del Centro de Investigaciones Biológicas (CSIC). Sobre el pGEM-T Easy con el homeobox I30E se introdujo la segunda mutación del doble mutante, V19R. Para las dos primeras PCRs del método de Higuchi *et al.* encaminadas a obtener esta segunda mutación se utilizaron en cada una 1 ng de plásmido pGEM-T Easy con el homeobox I30E, 100 pmol de cada cebador, 1mM dNTPs, 2mM MgCl_2 y 2,5 U Ecozyme, en un volumen de 100 μl . Para la segunda PCR se utilizaron 3,5 ng del fragmento situado en sentido 5' de la mutación, 1,6 ng del fragmento situado en sentido 3' de la mutación (ambos fragmentos de ADN generados en las primeras PCRs), 70 pmol de cada cebador, 1mM dNTPs, 2mM MgCl_2 y 3,5 U EcoTaq, en un volumen de 70 μl . El fragmento de ADN obtenido (homeobox de la variante **16**) fue ligado con el vector pGEM-T Easy, utilizando 13 ng de vector, 9 ng de inserto y 1,5 U de T4 DNA ligasa incluida con el kit del pGEM-T Easy, en un volumen de 5 μl . Nuevamente, con objeto de obtener cantidades manejables del plásmido pGEM-T Easy con el homeobox de **16**, la mezcla de ligación fue empleada para transformar células competentes de la cepa DH5 α de *Escherichia coli*. La correcta introducción de las mutaciones fue verificada por secuenciación en el Servicio de Secuenciación Automática de DNA del Centro de Investigaciones Biológicas (CSIC). Las condiciones de todas las PCRs del método están detalladas en la [Tabla 5.2](#). En todas las PCRs se introdujo un paso inicial de 94 °C 2 min, y un paso final de 72 °C 10 min. El mutante **16** en pGEM-T Easy fue digerido en primer lugar con la enzima de restricción BamHI (MBI Fermentas), usando 450 ng de pGEM-T Easy con el homeobox de **16**, y 25 U de BamHI, en un volumen de 50 μl , 1 hora a 37 °C. Tras purificar el ADN de la mezcla mediante extracción con fenol:cloroformo:alcohol isoamílico (Sigma), se llevó a cabo la digestión con NdeI (MBI Fermentas), utilizando 300 ng del digerido con BamHI y 12 U de NdeI, en un volumen de 34 μl , 1 hora a 37 °C. El homeobox del mutante **16** digerido así con BamHI y NdeI, fue ligado con el vector de expresión pT7-7 (cedido por el Dr. José Luis García, Centro de Investigaciones Biológicas, CSIC) digerido del mismo modo, usando 10 ng de vector, 5 ng de inserto y 5 U de T4 DNA ligasa (MBI Fermentas) en un volumen de 10 μl . Tal y como se ha descrito anteriormente, con objeto de obtener el plásmido pT7-7-**16**, se transformaron células DH5 α con la mezcla de ligación.

Los plásmidos pT7-7 con los homeobox del resto de mutantes se obtuvieron empleando el sistema QuikChange Site-Directed Mutagenesis Kit (Stratagene), mediante un único paso de PCR por cada sustitución individual, como se muestra en la [Figura 5.2](#). En base al protocolo descrito en el kit, se utilizó para cada reacción de mutagénesis 1 ng del correspondiente plásmido,

TABLA 5.2: Condiciones de PCR de las reacciones de mutagénesis dirigida

Mutante	Mutación	PCR	Nº de ciclos	Pasos de cada ciclo
1	P26L		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
2	E19D		18	95 °C 30 s, 50 °C 1 min, 62 °C 30 s, 68 °C 3,33 min
3	V15E		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
	K37R		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
4	V19R		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
5	V15R		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
	K37E		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
6	V15R		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
	K37E		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
7	V15R		16	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
8	I30R		16	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
9	K37E		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
10	K37E		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
11	V15E		16	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
12	K37R		16	95 °C 30 s, 55 °C 1 min, 62 °C 30 s, 68 °C 2,75 min
13	K37R		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
14	V19E		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
15	K37E		16	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
16	V19R	1 ^{as}	30	94 °C 1 min, 55 °C 50 s, 72 °C 1 min
		2 ^a	18	94 °C 1 min, 55 °C 50 s, 72 °C 50 s
	I30E	1 ^{as}	30	94 °C 1 min, 55 °C 50 s, 72 °C 40 s
		2 ^a	18	94 °C 1 min, 55 °C 50 s, 72 °C 50 s
17	P26L		18	95 °C 30 s, 55 °C 1 min, 59 °C 30 s, 68 °C 2,75 min
18	V15R		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
19	V19E		16	95 °C 30 s, 55 °C 1 min, 65 °C 30 s, 68 °C 2,5 min
20	K37E		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
21	K37L		18	95 °C 30 s, 50 °C 1 min, 62 °C 30 s, 68 °C 3,33 min
22	V15E		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
	K37R		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
23	V15E		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
24	V15E		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
25	I30E		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
26	K37R		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
27	E42D		18	95 °C 30 s, 50 °C 1 min, 62 °C 30 s, 68 °C 3,33 min
28	E42L		18	95 °C 30 s, 55 °C 1 min, 59 °C 30 s, 68 °C 2,75 min
29	R31L		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
30	L31I		18	95 °C 30 s, 50 °C 1 min, 62 °C 30 s, 68 °C 3,33 min
31	E42L		18	95 °C 30 s, 55 °C 1 min, 59 °C 30 s, 68 °C 2,75 min
32	Y17L		18	95 °C 30 s, 50 °C 1 min, 62 °C 30 s, 68 °C 3,33 min
33	E17Y		18	95 °C 30 s, 55 °C 1 min, 60 °C 30 s, 68 °C 2,75 min
	R52Y		18	95 °C 30 s, 55 °C 1 min, 59 °C 30 s, 68 °C 2,75 min

12,5 ng de cada cebador y 0,25 U de *PfuTurbo* DNA polimerasa, en un volumen de 5 µl. Los plásmidos de partida de cada mutante, así como los cebadores utilizados y el nombre de los plásmidos con los mutantes obtenidos, se especifican en la [Tabla 5.1](#). Todos los cebadores fueron encargados a Isogen

Life Science. Las condiciones de cada reacción de mutagénesis están detalladas en la [Tabla 5.2](#). En todas las reacciones se introdujo un paso inicial de 95 °C 30 s. Tras cada reacción de mutagénesis, se continuó con el protocolo recomendado mediante la digestión con la enzima de restricción DpnI incluida en el kit (1 U en los 5 µl resultantes de la PCR), 1 hora a 37 °C. Como último paso se transformaron células competentes de la cepa DH5α de *Escherichia coli*, que mostraron mayor eficiencia de transformación que las XL1-Blue incluidas en el kit, y que permitieron obtener cantidades manejables de cada plásmido mutante. La correcta introducción de las mutaciones fue verificada por secuenciación en el Servicio de Secuenciación Automática de DNA del Centro de Investigaciones Biológicas (CSIC), o bien en la Unidad de Genómica del Parque Científico de Madrid – UCM.

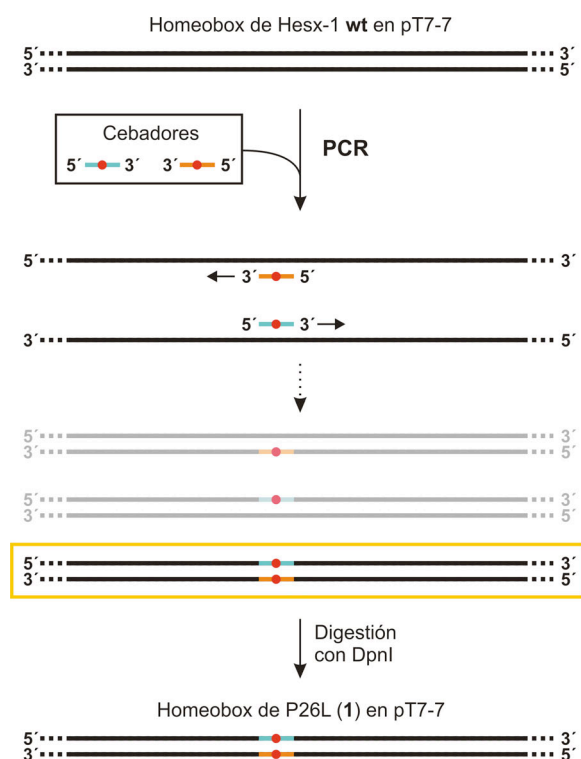


FIGURA 5.2: Esquema de funcionamiento del kit de mutagénesis dirigida QuikChange de Stratagene, que ilustra la introducción de la mutación simple P26L en la secuencia del homeobox salvaje, para obtener la variante 1. Los cebadores 1 y 2 utilizados se representan en turquesa y naranja, respectivamente (la numeración es la usada en la [Tabla 5.1](#)). La posición de las mutaciones se representa con un punto rojo. Las pequeñas líneas discontinuas a ambos lados del homeobox representan el plásmido en el que está insertado. Encuadrado en amarillo se muestra el principal producto final de la PCR. En tonos tenues se representan los productos minoritarios, que contienen ADN parental y son así susceptibles a la digestión final con DpnI. La flecha discontinua hace referencia a etapas del proceso de PCR omitidas para dar claridad.

5.1.2. Expresión y purificación proteica

Se transformaron células competentes de la cepa de *Escherichia coli* BL21 (DE3) (Stratagene) con cada uno de los plásmidos correspondientes a los diversos mutantes obtenidos. Con objeto de maximizar los niveles de expresión

del homeobox de Hesx-1 para cada mutante, se realizó un chequeo sistemático de distintas colonias. A su vez se optimizaron, en cada caso, los tiempos y temperaturas de expresión. De acuerdo con esta información, se destinaron cultivos a gran escala a la producción de cada homeodominio. Así, las células de las colonias seleccionadas se crecieron en 1 litro de medio LB (10 g/L triptona [Laboratorios Conda], 5 g/L extracto de levadura [Laboratorios Conda], 10 g/L NaCl [Panreac Química S.A.U.] pH 7,5, 100 µg/ml ampicilina (Sigma), hasta una densidad óptica a 600 nm ($DO_{600\text{ nm}}$) comprendida entre 0,7 y 0,8 a 37 °C. En ese momento, se indujo la expresión de los diferentes mutantes con 0,5 mM isopropil-β-D-tiogalactopiranosido (IPTG) (MBI Fermentas o Acros Organics). Los tiempos y temperaturas de inducción para cada mutante se recogen en la [Tabla 5.3](#).

El procedimiento empleado para la purificación de cada uno de los mutantes se detalla a continuación, y se representa esquemáticamente en la [Figura 5.3](#). En primer lugar, se recolectaron por centrifugación (3000 g, 15 min, 4 °C) las células de cada cultivo. Acto seguido se resuspendieron en 25 ml de tampón fosfato (Sigma-Aldrich) 50 mM pH 7,5, 400 mM NaCl (Panreac Química S.A.U.), 10% (p/v) sacarosa (Fisher Scientific o Merck), 2 mM β-mercaptoetanol (Aldrich), y se lisaron empleando una Prensa Francesa (Thermo Electron Corporation) a una presión de 1100 psi, previa adición de

TABLA 5.3: Condiciones de expresión para cada mutante en cultivos de BL21 (DE3)^a

Mutante	T _{expresión} (°C)	t _{expresión} (horas)	Mutante	T _{expresión} (°C)	t _{expresión} (horas)
wt	24/24	15,5/22	17	25/27/24	18/18/12,5
1	25/22	19/14,5	18	25/27	13,5/16
2	28	17	19	27/26,5	18/18
3	25	16	20	24	14
4	24/24/28	13,5/19/13	21	27	17
5	25/26,5	16/9	22	22	18
6	27	18	23	23/28	16,5/24
7	24/26	15,5/19	24	22	18
8	22/26,5	14,5/9	25	24	19
9	23	20	26	23/28	18/24
10	25/27	15/16	27	27	17
11	27/25	16,5/19	28	27	17
12	27/24	18/18	29	28	17
13	22	14,5	30	27	17
14	23	20	31	25/24/24/28/25	17/4/12,5/18/15
15	27/21	16,5/19	32	28	17
16	30/26	20,5/19,5	33	25/24	18/4

^a Los valores correspondientes a los distintos cultivos hechos para un mismo mutante se muestran separados por barras.

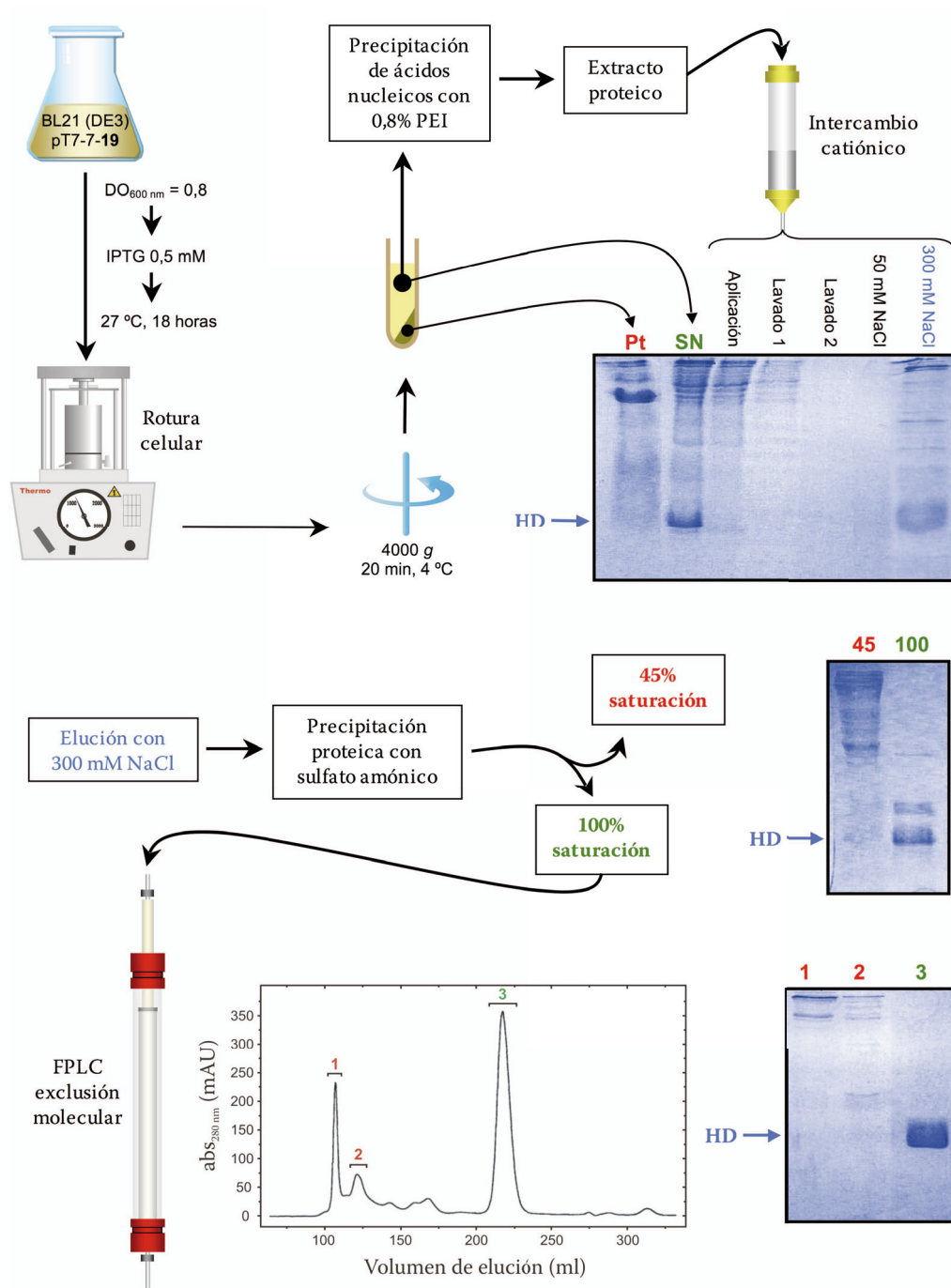


FIGURA 5.3: **Esquema del proceso de purificación** de las variantes del homeodominio de HESX-1, ejemplificado con el mutante 19.

1 mM de fluoruro de fenilmetilsulfonilo (PMSF) (Aldrich). A continuación, el lisado se centrifugó (4000 g, 20 min, 4 °C). El sobrenadante resultante se sometió a precipitación de ácidos nucleicos con 0,8% polietilenimina (Aldrich), 20 min a 0 °C, seguida de centrifugación (4000 g, 20 min, 4 °C). El nuevo sobrenadante se dializó una noche a 4 °C en 1 L de tampón fosfato 50 mM pH 7,5, 2 mM β -mercaptoetanol, usando membranas de diálisis Cellu-Sep T1, MWCO 3500 (Orange Scientific). Tras centrifugar (16000 g, 20 min, 4 °C), el dializado se aplicó a una columna de intercambio catiónico Bio-Rex 70 (Bio-Rad), con 25 ml de resina empaquetada, previamente equilibrada en tampón fosfato 50 mM pH 7,5, 2 mM β -mercaptoetanol, 0,01% azida sódica (Sigma-Aldrich) (tampón A). Tras un lavado con 100 ml de tampón A, y otro con 50 ml de tampón A con 50 mM NaCl, se procedió a la elución con 50 ml de tampón A con 300 mM NaCl. Estos volúmenes y concentraciones de NaCl son el resultado de una optimización previa mediante la aplicación de un gradiente de NaCl. Para regenerar la columna tras cada uso se le pasaron 200 ml de tampón A con 1 M NaCl. El flujo durante todo el proceso fue de 1 ml/min. A continuación, el eluido se sometió a fraccionamiento con sulfato amónico (Panreac Química S.A.U.). Para ello, se añadió sulfato amónico hasta un 45% de saturación a 0 °C, seguido de centrifugación (12000 g, 20 min, 4 °C). Al sobrenadante resultante se le añadió sulfato amónico hasta un 100% de saturación, nuevamente seguido de centrifugación (12000 g, 20 min, 4 °C). Finalmente, se realizó un paso de purificación empleando FPLC. Con este objetivo, el precipitado de la última centrifugación se redisolvió en 5 ml de tampón fosfato 20 mM pH 7,0, 100 mM NaCl, 2 mM β -mercaptoetanol, 0,01% azida sódica (tampón B). Tras centrifugar (16000 g, 5 min, T_{amb}) para separar los restos insolubles, el sobrenadante se inyectó en una columna cromatográfica de exclusión molecular Superdex-75 HiLoad 26/60 (Amersham Biosciences), previamente equilibrada en tampón B, y se pasaron 2 volúmenes de tampón B, a un flujo de 2,5 ml/min. Un perfil típico de esta purificación cromatográfica se muestra en la parte inferior de la [Figura 5.3](#). Las fracciones correspondientes al homeodominio se juntaron para dar un volumen de unos 22 ml, con una pureza estimada en más del 95% por SDS-PAGE. Asimismo, mediante espectrometría de masas MALDI-TOF (Centro de Investigaciones Biológicas, CSIC) se comprobó que el peso molecular fuese el esperado. La disolución final se desalinizó mediante diálisis durante una noche a 4 °C en 1 L de tampón fosfato 2 mM, pH 6,0, 5 mM NaCl, 2 mM β -mercaptoetanol, y a continuación se liofilizó. Los liofilizados se conservaron a -80 °C.

5.2

EXPERIMENTOS DE DICROÍSMO CIRCULAR

Los experimentos de dicroísmo circular se realizaron en un espectropolarímetro JASCO J-810, equipado con un accesorio de control de temperatura Peltier JASCO PTC-423S.

5.2.1. Curvas de desnaturalización química

Las muestras empleadas para la obtención de las curvas de desnaturalización química se prepararon disolviendo liofilizado de proteína en tampón fosfato 10 mM pH 6,0, 0,5 mM DL-ditiotreitol (DTT) (Sigma), 100 mM NaCl (tampón C), y dializando frente al mismo tampón a 4 °C durante una noche. Finalmente, las muestras se centrifugaron (16000 g, 10 min, 4 °C) con objeto de eliminar agregados insolubles.

Se realizaron experimentos de desnaturalización química de cada polipéptido a 20 °C, empleando cubetas Hellma de 1 cm de paso óptico y 2 ml de volumen, con agitación magnética de 850 rpm. La concentración de proteína fue en todos los casos de en torno a 18 μ M. Los perfiles de desnaturalización se obtuvieron monitorizando las variaciones en la elipticidad a 222 nm al incrementar la concentración de desnaturalizante (urea [Fluka] o cloruro de guanidinio [GdmCl] [Pierce]) mediante sucesivas adiciones de una solución concentrada (urea 9,45 M, o GdmCl 8,36 M) en tampón C. Para ello se utilizó un sistema de inyección automática (JASCO ATS-429S). Después de cada adición, se permitió un periodo de equilibrado de 1 minuto, tras el cual se registró cada uno de los puntos mediante 3 acumulaciones. El resto de parámetros del espectropolarímetro se ajustaron del siguiente modo: Band width: 2 nm; Response: 4 s; Sensitivity: Standard. Las concentraciones molares de desnaturalizante, tanto de las soluciones concentradas, como de las muestras inicial y final de cada experimento, se confirmaron mediante medidas de índice de refracción, empleando las siguientes relaciones (Pace, 1986):

$$[\text{urea}] = 117,66 \cdot (\Delta N) + 29,753 \cdot (\Delta N)^2 + 185,56 \cdot (\Delta N)^3 \quad (1)$$

$$[\text{GdmCl}] = 57,147 \cdot (\Delta N) + 38,68 \cdot (\Delta N)^2 - 91,60 \cdot (\Delta N)^3 \quad (2)$$

En ellas, ΔN es la diferencia entre el índice de refracción de la solución con y sin agente desnaturizante.

Para algunos de los mutantes se registraron curvas de renaturalización, cuya concordancia con las de desnaturalización confirmaron la reversibilidad del proceso.

Para el análisis de los perfiles de desnaturalización se asumió un modelo de dos estados, ajustándose los datos experimentales a la siguiente ecuación (Santoro y Bolen, 1988), utilizando el programa Origin 6.0:

$$y = \frac{(y_N + m_N[D]) + (y_U + m_U[D]) \cdot \exp - \frac{\Delta G_{U(H_2O)} - m[D]}{RT}}{1 + \exp - \frac{\Delta G_{U(H_2O)} - m[D]}{RT}} \quad (3)$$

En esta expresión, R es la constante de los gases, T la temperatura en grados Kelvin, y la elipticidad observada, $[D]$ la concentración del agente desnaturizante, y_N e y_U los valores de y en la intersección a $[D] = 0$ de las líneas base previa y posterior a la transición, respectivamente, m_N y m_U las pendientes de dichas líneas base, m la dependencia de ΔG_U (variación de energía libre asociada al proceso de desnaturalización) con la concentración de desnaturizante, y $\Delta G_{U(H_2O)}$ el valor de ΔG_U en ausencia de agente desnaturizante. Los errores estándar de los valores $\Delta G_{U(H_2O)}$ y m derivados de los ajustes de cada una de las curvas fueron, en todos los casos, inferiores a 0,15 kcal/mol y 0,03 kcal/(mol·M), respectivamente. Debido al criterio de signos utilizado en esta tesis, los valores de ΔG_F y m que se dan en “Resultados” son los opuestos de los valores de $\Delta G_{U(H_2O)}$ y m obtenidos de los ajustes.

5.2.2. Curvas de desnaturalización térmica

Para las curvas de desnaturalización térmica, las muestras se prepararon disolviendo liofilizado de proteína en tampón C, pero con NaCl a la concentración deseada para el experimento (0, 0,1, 0,4, 1,0 ó 2,0 M), y dializando frente al mismo tampón una noche a 4 °C. Finalmente, el dializado se centrifugó (16000 g, 10 min, 4 °C) para eliminar agregados insolubles.

Las curvas de desnaturalización térmica se realizaron, por tanto, a distintas concentraciones de NaCl (0, 0,1, 0,4, 1,0 y 2,0 M), en cubetas de cuarzo

Hellma de 1 cm de paso óptico y 1 ml de volumen. La concentración de proteína fue en todos los casos de en torno a 3 μM . Las curvas se adquirieron incrementando la temperatura a un ritmo constante de 20 grados/hora, y monitorizando los cambios de elipticidad a 222 nm. El resto de parámetros del espectropolarímetro se ajustaron del siguiente modo: Band width: 2 nm; Response: 4 s; Sensitivity: Standard. Tras cada experimento, se retornó a las condiciones de temperatura iniciales, comprobándose para todas las proteínas una recuperación de la señal inicial superior al 90%, lo que confirmó la reversibilidad del proceso de desnaturalización. Los datos experimentales se analizaron asumiendo un modelo de dos estados, mediante su ajuste a la siguiente expresión (Mayr *et al.*, 1993), utilizando el programa Origin 6.0 (Microcal Software):

$$y = \frac{(y_N + m_N T) - (y_U + m_U T)}{1 + \exp \left[-\frac{\Delta H_{U(T_m)}}{R} \left(\frac{1}{T} - \frac{1}{T_m} \right) - \frac{\Delta C_p}{R} \left(1 - \frac{T_m}{T} + \ln \frac{T_m}{T} \right) \right]} + y_U + m_U T \quad (4)$$

En ella, R representa la constante de los gases, y es la elipticidad observada, T la temperatura en grados Kelvin, T_m la temperatura de desnaturalización, es decir, el valor de T a mitad de la transición, y_N y y_U los valores de y en la intersección a $T = 0$ K de las líneas base previa y posterior a la transición, respectivamente, m_N y m_U las pendientes de dichas líneas base, $\Delta H_{U(T_m)}$ la entalpía van't Hoff asociada al proceso de desnaturalización a $T = T_m$, y ΔC_p la diferencia de capacidad calorífica entre el estado desnaturalizado y el nativo. Debido al criterio de signos utilizado en esta tesis, los valores de $\Delta H_{U(T_m)}$ obtenidos de los ajustes se cambiaron de signo para obtener la variación de entalpía asociada al proceso de plegamiento, $\Delta H_{F(T_m)}$. Cabe mencionar que algunas curvas, correspondientes a ciertas variantes termoestables, no completaron su transición a los 100 °C de temperatura. Para conseguir una buena aproximación de los ajustes en estos casos, se asignaron unos valores de y_U y m_U basados en los derivados para un número de curvas completas de variantes de composición similar.

A partir de los valores de T_m y $\Delta H_{F(T_m)}$ derivados del análisis de las curvas, y asumiendo un ΔC_p asociado al proceso de plegamiento de $-0,7$ kcal/(mol·K) para todos los mutantes, se obtuvieron los valores de $\Delta G_F(T)$ (variación de energía libre asociada al proceso de plegamiento a una temperatura determinada) aplicando la ecuación de Gibbs-Helmholtz:

$$\Delta G_F(T) = \Delta H_{F(T_m)} \left(1 - \frac{T}{T_m} \right) - \Delta C_p \left(T_m - T + T \ln \frac{T}{T_m} \right) \quad (5)$$

La temperatura utilizada para sustituir en la ecuación de Gibbs-Helmholtz fue de 77 °C (350 K) para el análisis comparativo de la proteína salvaje con los mutantes con mayor T_m (28, 29, 30 y 31) y de 55 °C (328 K) para el correspondiente al resto de mutantes (ver [Tabla 3.1](#)). Ambas temperaturas son la media de las T_m de todos los mutantes de cada grupo a todas las concentraciones de NaCl, de modo que las extrapolaciones que hubo que realizar fueron en global lo más pequeñas posible.

El valor de $\Delta C_p = -0,7$ kcal/(mol·K) representa la media de los valores obtenidos en los análisis de cada una de las curvas de desnaturalización térmica. Además, se corresponde bien con la disminución del área de superficie expuesta estimada para el plegamiento del homeodominio (Myers *et al.*, 1995). El hecho de que los valores m exhiban muy pequeñas variaciones para los diferentes mutantes (en todos los casos están en el rango de 0,8 a 1,0 kcal/[mol·M] para los experimentos de desnaturalización realizados con urea, y de 1,5 a 1,7 kcal/[mol·M] para los realizados con GdmCl) apoya el uso de un ΔC_p común para todos ellos (Myers *et al.*, 1995). Se han descrito con anterioridad valores de ΔC_p comparables para los homeodominios de Engrailed ($-0,7$ kcal/[mol·M] [Stollar *et al.*, 2003]) y de vnd (NK-2) ($-0,52$ kcal/[mol·M] [Gonzalez *et al.*, 2001]). De todos modos, no es de esperar que la falta de precisión que pudiera tener el valor de ΔC_p afecte a los resultados obtenidos, ya que todos los mutantes muestran T_m s relativamente cercanas a 55 ó 77 °C (dependiendo del grupo de mutantes), con lo que no fue necesario efectuar grandes extrapolaciones para determinar ΔG_F a esas temperaturas.

Los errores estándar correspondientes a T_m y $\Delta H_{F(T_m)}$ derivados de los ajustes de cada una de las curvas de desnaturalización térmica, fueron típicamente inferiores a 0,1 °C y 1 kcal/mol, respectivamente. Evidentemente, el error en los valores de ΔG_F a 55 ó 77 °C calculados mediante la ecuación de Gibbs-Helmholtz dependerá de los correspondientes a los valores de T_m y $\Delta H_{F(T_m)}$, y será tanto mayor cuanto mayor sea la diferencia entre la temperatura elegida (55 ó 77 °C en nuestro caso) y T_m . Para estimar los errores máximos, se calcularon los valores de ΔG_F considerando una variación de 2 kcal/mol en $\Delta H_{F(T_m)}$ ($\Delta H_{F(T_m)} \pm 2$ kcal/mol), 0,2 °C en T_m ($T_m \pm 0,2$ °C), y 0,2 kcal/(mol·K) en ΔC_p ($\Delta C_p \pm 0,2$ kcal/[mol·K]). La máxima diferencia entre las energías libres de plegamiento derivadas de esta forma se tomó como el error de ΔG_F para cada uno de los mutantes (ver la [Figura 5.4](#)).

Finalmente, siguiendo la metodología descrita por Schmid y colaboradores (Perl *et al.*, 2000; Martin *et al.*, 2002; Wunderlich *et al.*, 2005; Max

et al., 2007), se asumió que en el ΔG_F medido a 0 M NaCl están presentes los aportes procedentes de interacciones tanto electrostáticas como no polares ($\Delta G_{\text{total}} = \Delta G_F \text{ 0 M NaCl}$). Las contribuciones a la estabilidad todavía presentes a 2 M NaCl reflejan primordialmente interacciones de empaquetamiento e hidrófobas ($\Delta G_{\text{no polar}} = \Delta G_F \text{ 2 M NaCl}$), mientras que las contribuciones que serían apantalladas al añadir 2 M NaCl se atribuyeron a interacciones coulombicas ($\Delta G_{\text{coulomb}}$). Los términos $\Delta G_{\text{coulomb}}$ fueron calculados a partir de la diferencia $\Delta G_{\text{coulomb}} = \Delta G_F \text{ 0 M NaCl} - \Delta G_F \text{ 2 M NaCl}$, y los correspondientes errores (E) estimados como $E_{\Delta G_{\text{coulomb}}} = [(E_{\Delta G_F \text{ 0 M NaCl}})^2 + (E_{\Delta G_F \text{ 2 M NaCl}})^2]^{1/2}$.

Es conveniente señalar que esta asunción representa una simplificación significativa. De hecho, dada la localización parcialmente expuesta en la estructura 3D de HESX-1 de los residuos que vamos a tratar, la desolvatación parcial desfavorable de cadenas laterales cargadas (E o R) en cualquiera de estos lugares podría contribuir a $\Delta G_{\text{no polar}}$. Además, esta aproximación no tendría

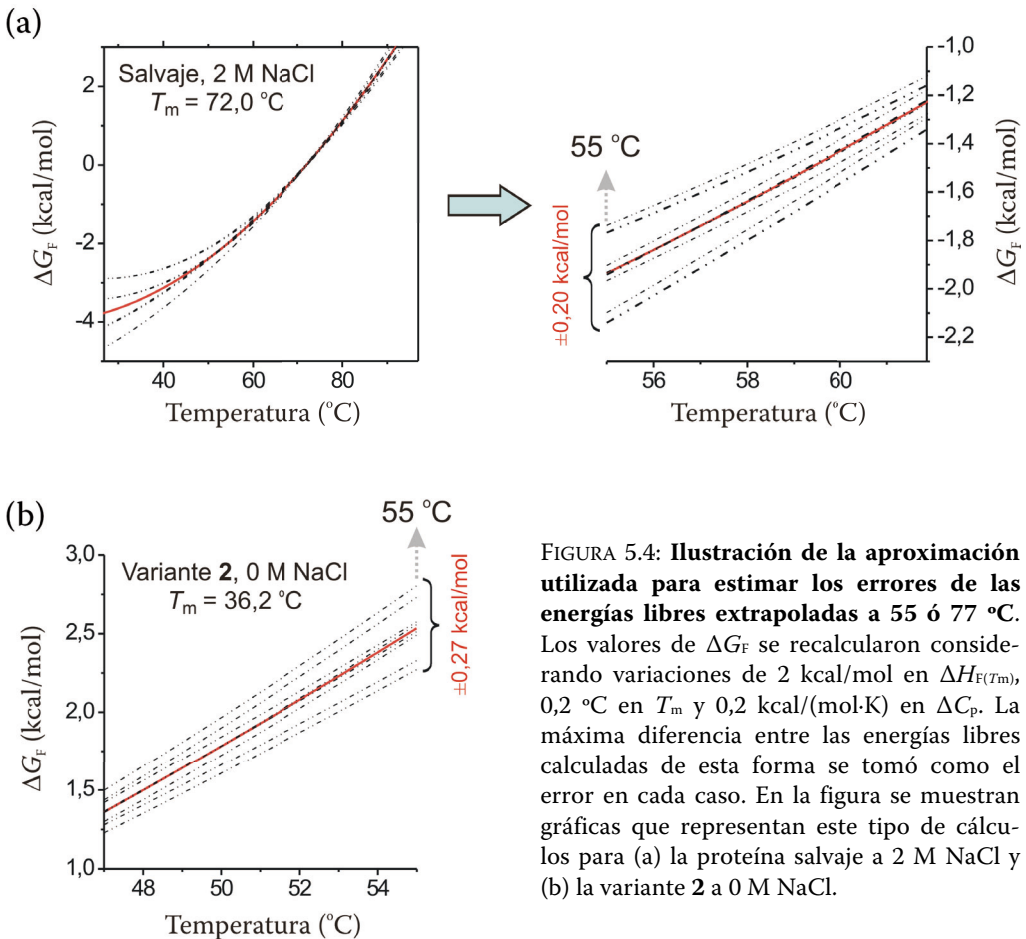


FIGURA 5.4: Ilustración de la aproximación utilizada para estimar los errores de las energías libres extrapoladas a 55 ó 77 $^\circ\text{C}$. Los valores de ΔG_F se recalcularon considerando variaciones de 2 kcal/mol en $\Delta H_F(T_m)$, 0,2 $^\circ\text{C}$ en T_m y 0,2 kcal/(mol·K) en ΔC_p . La máxima diferencia entre las energías libres calculadas de esta forma se tomó como el error en cada caso. En la figura se muestran gráficas que representan este tipo de cálculos para (a) la proteína salvaje a 2 M NaCl y (b) la variante 2 a 0 M NaCl.

en cuenta contribuciones coulómbicas que pudieran persistir a 2 M NaCl (Lindman *et al.*, 2006), por lo que $\Delta G_{\text{coulomb}}$ podría representar un límite inferior de las contribuciones electrostáticas a la estabilidad de la proteína (Mueller *et al.*, 2000; Wunderlich *et al.*, 2005). En relación a esto, numerosos estudios han demostrado que el efecto de apantallado que la fuerza iónica produce en las interacciones carga-carga aumenta con la distancia entre los residuos involucrados y su grado de exposición al disolvente (Kao *et al.*, 2000; Takano *et al.*, 2000; Dominy *et al.*, 2002; Lee *et al.*, 2002a, b). En el caso de los homeodominios, el carácter altamente dinámico y parcialmente expuesto de los pares cargados sugiere que sus interacciones deben verse significativamente atenuadas por la sal, con lo que la aproximación empleada parece razonable.

5.3

EXPERIMENTOS DE RESONANCIA MAGNÉTICA NUCLEAR (RMN)

5.3.1. Experimentos de RMN monodimensionales de los homeodominios salvaje, 31, 17 y 19

Una estimación adicional de la estabilidad térmica relativa de las proteínas salvaje y **31** se obtuvo mediante experimentos de desnaturalización que se siguieron empleando RMN. Con este objetivo se prepararon muestras de ambos polipéptidos (a una concentración aproximada de 200 μ M) en tampón fosfato 10 mM pH 6,0, 2 mM DTT, 100 mM NaCl, en mezcla 85:15 H₂O:D₂O. En un segundo paso, se adquirieron espectros monodimensionales en un espectrómetro Bruker Avance 800, a temperaturas comprendidas entre 20 y 70 °C. En todos los casos, la supresión de agua se consiguió mediante la utilización de un módulo Watergate (Piotto *et al.*, 1992). Se adquirieron 48 acumulaciones por espectro con tiempos de adquisición y de relajación de 1 segundo.

Además de estos experimentos, adquirimos espectros 1D adicionales de los mutantes **17** y **19**, con objeto de estimar la magnitud de las alteraciones estructurales promovidas por las mutaciones. Para ello utilizamos unos procedimientos análogos a los arriba descritos, y una temperatura de 20 °C.

5.3.2. Determinación de la estructura tridimensional de la variante termoestable 31 (R31L/E42L) por RMN

Para el análisis estructural del mutante **31** por RMN se prepararon muestras en tampón fosfato 10 mM pH 6,0, 2 mM DTT, 100 mM NaCl, tanto en D₂O como en mezcla 85:15 H₂O:D₂O. La concentración de proteína empleada (de acuerdo con su absorbancia a 280 nm) estuvo, en todos los casos, en el rango de 400-800 μ M.

La adquisición de datos se llevó a cabo utilizando espectrómetros Bruker Avance 800, y Bruker Avance 600 (equipado con criosonda). Para cada una de

las muestras se realizaron experimentos 2D NOESY (Kumar *et al.*, 1980), 2D TOCSY (Bax y Davis, 1985) y DQF-COSY (Marion y Wüthrich, 1983) a 25 °C, que se llevaron a cabo en el modo sensible a la fase (Redfield y Kuntz, 1975; States *et al.*, 1982; Marion y Wüthrich, 1983) usando el método TPPI (Marion y Wüthrich, 1983) para la detección en cuadratura en F_1 . En los experimentos en H_2O la supresión de agua se consiguió mediante la utilización de un módulo Watergate (Piotto *et al.*, 1992). Se usaron matrices de datos de 512×2048 puntos para digitalizar una anchura espectral de 8000-6000 Hz en ambas dimensiones. Se adquirieron 80 acumulaciones por incremento con un tiempo de espera de 1 s. Antes de aplicar la transformada de Fourier, se utilizó llenado con ceros en F_1 , para expandir los datos a una matriz de 1024×2048 puntos. Finalmente, se corrigió la línea base en F_1 y F_2 .

Los experimentos TOCSY se realizaron utilizando una secuencia MLEV-17 (Bax y Davis, 1985) durante el tiempo de mezcla isotrópica, que fue de 60 ms. Los experimentos NOESY se llevaron a cabo con tiempos de mezcla de 75 y 150 ms.

El proceso de asignación de las resonancias de protón de **31** se llevó a cabo, en un primer paso, mediante la comparación de sus espectros con los correspondientes al polipéptido salvaje. La lista inicial obtenida de esta forma se completó siguiendo el método estándar desarrollado por Wüthrich y colaboradores (Wüthrich, 1986), basado en el análisis de los espectros 2D NOESY, 2D TOCSY y DQF-COSY.

Los límites superiores para las distancias protón-protón se obtuvieron a partir de las intensidades de los picos de cruce de los NOESY adquiridos con tiempos de mezcla de 75 y 150 ms. Éstos fueron clasificados, de acuerdo con su integral, como intenso, medio y débil, correspondiendo a límites superiores de 2,5, 3,5 y 5,0 Å, respectivamente. El límite inferior para las distancias protón-protón se estableció como la suma de los radios van der Waals de los mismos. Para los grupos metilo presentes en alaninas, valinas, leucinas, isoleucinas y treoninas se utilizaron pseudoátomos, incrementándose el límite superior de la restricción en 1 Å. De un modo similar, para los anillos aromáticos de fenilalanina y tirosina las restricciones se aplicaron sobre los carbonos C_γ y C_ζ , incrementándose el límite superior en 2,2 Å.

Los cálculos de la estructura de **31** fueron realizados en un ordenador O2 Silicon Graphics, empleando el programa DYANA (Guntert *et al.*, 1997). En la ronda final se obtuvieron 400 estructuras empleando un conjunto de 703 restricciones de distancias. Las 30 mejores, en términos de función diana y de violaciones, fueron sometidas a un refinamiento posterior empleando un protocolo de enfriamiento simulado (*simulated annealing*) (Scheek *et al.*, 1989) con el paquete de programas AMBER 5.0 y los parámetros descritos por Kollman y colaboradores (Cornell *et al.*, 1995). Con objeto de evitar interac-

ciones no realistas entre regiones desordenadas de la proteína y el núcleo helicoidal estructurado, se utilizaron en estos cálculos moléculas de agua explícitas TIP3P (Cornell *et al.*, 1995) y condiciones periódicas de contorno. Después de un paso inicial de minimización restringida de energía (REM), con 1000 iteraciones del algoritmo “steepest-descent”, las estructuras fueron calentadas a 500 K en 10 ps. A esta temperatura, se llevaron a cabo 10 ps de dinámica molecular restringida (RMD) empleando un tiempo de integración de 1 fs. Las estructuras se sometieron entonces a un enfriamiento durante 20 ps hasta alcanzar una temperatura de 1 K. Las estructuras finales fueron sometidas a minimización restringida de energía (REM), utilizando 2000 iteraciones del algoritmo de gradientes conjugados.

5.3.3. Medidas de pK_a en los mutantes cuádruples 12 y 15 por RMN

Con objeto de determinar los valores de pK_a correspondientes a los residuos ácidos presentes en los mutantes **12** y **15**, se prepararon muestras de RMN en tampón fosfato 10 mM, 2 mM DTT, 100 mM NaCl, en mezcla 85:15 H₂O:D₂O. La concentración de proteína empleada (de acuerdo con su absorbancia a 280 nm) estuvo, en todos los casos, en el rango de 400-600 μ M. A continuación se adquirieron espectros 2D NOESY y 2D TOCSY a 5 °C, a diferentes valores de pH entre 2 y 7 (ajustando, en cada caso, mediante la adición de pequeñas alícuotas de HCl o NaOH).

La adquisición de datos se llevó a cabo utilizando un espectrómetro Bruker Avance 800 utilizando métodos análogos a los anteriormente descritos para la elucidación estructural de **31**. Las resonancias de los protones de todas las cadenas laterales de ácido aspártico y glutámico de las variantes **12** y **15** fueron asignadas a los distintos pHs, tomando como punto de partida las asignaciones ya conocidas para la proteína salvaje y siguiendo el método estándar.

A continuación, monitorizamos los cambios en los desplazamientos químicos de los protones de cadena lateral H β y H γ , en ácidos glutámicos, y H β en ácidos aspárticos, en función del pH. Las curvas de valoración así obtenidas se analizaron con el programa Origin 6.0 para derivar los valores de pK_a . Con este objeto llevamos a cabo un ajuste no lineal de mínimos cuadrados empleando una forma modificada de la ecuación de Henderson-Hasselbalch:

$$\delta(\text{pH}) = \frac{\delta_a + \delta_b \cdot 10^{n(\text{pH} - pK_a)}}{1 + 10^{n(\text{pH} - pK_a)}} \quad (6)$$

En esta expresión, $\delta(\text{pH})$ es el desplazamiento químico en función del pH, δ_a y δ_b los valores de δ en los extremos de pH ácido y básico, respectivamente, y n el coeficiente de Hill, que se fijó en 1.

Para obtener la diferencia de estabilidad entre la proteína con el grupo i (grupo carboxilo en cuestión) completamente protonado y con el grupo i completamente desprotonado, utilizamos la siguiente ecuación:

$$\Delta\Delta G_F^i = 2,3RT(pK_a^{F,i} - pK_a^{U,i}) \quad (7)$$

En ella, R representa la constante de los gases, T la temperatura en grados Kelvin, y $pK_a^{F,i}$ y $pK_a^{U,i}$ son los pK_a del grupo i en la proteína plegada y desplegada, respectivamente. El valor de pK_a calculado mediante la ecuación (6) se corresponde con $pK_a^{F,i}$. Los valores de $pK_a^{U,i}$ utilizados fueron 3,9 para ácido aspártico y 4,3 para ácido glutámico (Nozaki y Tanford, 1967; Yang y Honig, 1993). Si el grupo desprotonado estabiliza el estado nativo, $pK_a^{F,i}$ será menor que $pK_a^{U,i}$, y entonces $\Delta\Delta G_F^i$ será negativo. En cambio, si el grupo protonado estabiliza el estado nativo, $pK_a^{F,i}$ será mayor que $pK_a^{U,i}$, y $\Delta\Delta G_F^i$ será positivo.

5.4

EXPERIMENTOS DE UNIÓN A ADN POR MICROCALORIMETRÍA

Se realizaron estudios de asociación homeodominio-ADN a 15, 25 y 35 °C, en tampón fosfato 20 mM pH 6,0, 150 mM NaCl, 5 mM MgCl₂ (Acros Organics) (tampón D), empleando un microcalorímetro VP-ITC (MicroCal, LLC), con una celda de reacción de 1,467 ml. Oligonucleótidos de 14 bases (5'-GTCTAATTGACGCG-3' y su complementaria 5'-CGCGTCAATTAGAC-3'), que incorporan una única secuencia 5'-TAATTG-3' de unión específica de HESX-1, fueron adquiridos de Isogen Life Science. Ambas hebras se hibridaron para formar el dúplex de ADN correspondiente. Con este objetivo, se mezclaron cantidades equimoleculares de las mismas previamente disueltas en tampón D, y se calentó la mezcla a 90 °C 10 min, seguido de enfriamiento lento hasta alcanzar la temperatura ambiente. Por otro lado, se disolvió liofilizado de proteína en tampón D. A continuación, tanto la proteína como el ADN fueron dializados exhaustivamente frente a tampón D. Por norma general, se valoraron soluciones 4,0-14,0 µM de dúplex de ADN con soluciones 100-200 µM de las distintas variantes del homeodominio contenidas en una jeringa de 300 µl. Al menos 30 inyecciones consecutivas de 5 a 10 µl fueron aplicadas a intervalos de 5 min, manteniendo una agitación constante de 300 rpm en la celda. Como control, se realizaron experimentos de dilución de los polipéptidos en tampón D. Los calores de dilución de la proteína en la solución de ADN estuvieron de acuerdo con los obtenidos al inyectar las proteínas en el mismo volumen de tampón, y fueron restados a los correspondientes al proceso de asociación.

Los datos así obtenidos fueron procesados empleando el *software* suministrado con el instrumento por MicroCal, LLC, teniendo en cuenta las concentraciones de proteína y ADN, y los volúmenes aplicados y desplazados en cada inyección. Las curvas de valoración obtenidas de esta forma fueron analizadas con el programa Origin 7.0, también suministrado con el aparato, usando para el ajuste un modelo de unión de n sitios equivalentes. Para ello se utiliza la siguiente ecuación:

$$Q(i) = \frac{n[A]_{\text{tot}} \Delta H_b V_{\text{cel}}}{2} \left[1 + \frac{[P]_{\text{tot},i}}{n[A]_{\text{tot}}} + \frac{1}{nK_b[A]_{\text{tot}}} - \sqrt{\left(1 + \frac{[P]_{\text{tot},i}}{n[A]_{\text{tot}}} + \frac{1}{nK_b[A]_{\text{tot}}} \right)^2 - \frac{4[P]_{\text{tot},i}}{n[A]_{\text{tot}}}} \right] \quad (8)$$

En esta expresión, Q es el calor total desprendido tras la inyección i , n el número de sitios de unión idénticos e independientes de la proteína al ADN, ΔH_b la variación de entalpía asociada a la unión, V_{cel} el volumen de la celda de reacción, $[P]_{\text{tot},i}$ la concentración total de proteína añadida a la celda tras la inyección i , $[A]_{\text{tot}}$ la concentración total de ADN en la celda, y K_b la constante de unión. El ajuste proporciona los valores de n , ΔH_b y K_b . A partir de K_b , se obtuvo la energía libre de asociación (ΔG_b) utilizando la siguiente ecuación:

$$\Delta G_b = -RT \cdot \ln K_b \quad (9)$$

En ella, R representa la constante de los gases, y T la temperatura en grados Kelvin.

Para cada una de las variantes del homeodominio, los parámetros termodinámicos derivados de esta manera fueron el promedio de dos experimentos independientes.

CONCLUSIONES

1. Nuestro análisis de secuencias de la base de datos “Homeodomain Resource” pone de manifiesto la presencia, en una fracción significativa de homeodominios, de tres puentes salinos interhélice que conectan las posiciones 19-30, 31-42 y 17-52. Por el contrario, la presencia de pares hidrófobos en dichas localizaciones es altamente infrecuente.

2. Nuestros resultados experimentales indican que la estabilización aportada por pares alifáticos en las posiciones fronterizas 19-30 y 31-42 es significativamente mayor que la proporcionada por puentes salinos en idénticas localizaciones. En el caso del par 19-30, dicho comportamiento se mantiene, con independencia de la polaridad del puente salino, incluso en aquellas situaciones en las que es posible la formación de una red cooperativa que involucre la tétrada 19, 30, 15 y 37. Los experimentos a distintas concentraciones de sal permiten sugerir que la mayor estabilización por residuos alifáticos refleja, principalmente, una mejora en las interacciones de empaquetamiento con el núcleo hidrófobo.

3. La combinación L31/L42, presente en la variante **31**, genera una versión termoestable del homeodominio de HESX-1. Su estructura tridimensional, resuelta por métodos de RMN, pone de manifiesto una reorganización de carácter local del núcleo hidrófobo. Esta observación es consistente con el estudio de estabilidad descrito en la presente memoria.

4. A pesar de tener un efecto menos favorable que los pares alifáticos, los puentes salinos en posiciones fronterizas 19-30 y 31-42 contribuyen significativamente a la estabilidad del homeodominio, mediante el empaquetamiento de sus cadenas hidrocarbonadas con el núcleo hidrófobo de la proteína. La desestabilización inducida por la eliminación de grupos metileno mediante mutaciones E→D en las variantes **2** y **27** ilustra este efecto.

5. A diferencia del comportamiento observado para los pares 19-30 y 31-42, las posiciones 17-52 sólo parecen tolerar residuos polares. Su alto grado de exposición al disolvente podría estar en el origen de los problemas de solubilidad y estabilidad encontrados para las variantes con pares hidrófobos en estas posiciones.

6. De acuerdo con un estudio estadístico de secuencias previo, las posiciones 26 y 19-30 están altamente correlacionadas. De esta forma, cuando 19-30 porta un par alifático, 26 es siempre prolina, y cuando 19-30 está ocupado por un puente salino, 26 tiende a ser alifático. Nuestros resultados experimentales concuerdan con estas tendencias, y destacan la inviabilidad de la combinación V19/L26/I30 presente en la variante 1.

7. Nuestros ensayos de unión específica a ADN confirman la participación de la arginina 31 en el reconocimiento. No obstante, su aporte a la estabilización del complejo es relativamente modesto en comparación con el descrito para interacciones similares en otros complejos proteína/ADN. Por su parte, la influencia del ácido glutámico 42, o de su puente salino con R31, es prácticamente despreciable en este aspecto.

8. El marcado efecto de las posiciones 19, 30, 15 y 37 en la estabilidad del homeodominio contrasta con una influencia prácticamente nula de estos residuos en la afinidad de unión a ADN. Este resultado pone de manifiesto una llamativa ausencia de correlación entre estabilidad y función. Estudios previos han demostrado que el ajuste inducido de las regiones desestructuradas del homeodominio juega un papel clave en el proceso de reconocimiento. Por el contrario, de acuerdo con nuestros datos, el potencial ajuste conformacional del núcleo estructurado parece tener un efecto despreciable sobre la estabilidad del complejo. Estas propiedades permitirían una modulación independiente de estabilidad y unión a ADN por parte de la evolución.

REFERENCIAS

- Ades, S.E., Sauer, R.T., 1994. Differential DNA-binding specificity of the engrailed homeodomain: the role of residue 50. *Biochemistry* **33**, 9187-9194.
- Affolter, M., Percival-Smith, A., Müller, M., Leupin, W., Gehring, W.J., 1990. DNA binding properties of the purified Antennapedia homeodomain. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 4093-4097.
- Amendt, B.A., Sutherland, L.B., Russo, A.F., 1999. Multifunctional role of the Pitx2 homeodomain protein C-terminal tail. *Molecular and Cellular Biology* **19**, 7001-7010.
- Banerjee-Basu, S., Ryan, J.F., Baxevanis, A.D., 2000. The homeodomain resource: a prototype database for a large protein family. *Nucleic Acids Research* **28**, 329-330.
- Banerjee-Basu, S., Sink, D.W., Baxevanis, A.D., 2001. The homeodomain resource: sequences, structures, DNA binding sites and genomic information. *Nucleic Acids Research* **29**, 291-293.
- Banerjee-Basu, S., Moreland, T., Hsu, B.J., Trout, K.L., Baxevanis, A.D., 2003. The Homeodomain Resource: 2003 update. *Nucleic Acids Research* **31**, 304-306.
- Barril, X., Alemán, C., Orozco, M., Luque, F.J., 1998. Salt bridge interactions: stability of the ionic and neutral complexes in the gas phase, in solution, and in proteins. *Proteins: Structure, Function and Genetics* **32**, 67-79.
- Bax, A., Davis, D.G., 1985. MLEV-17 based two-dimensional homonuclear magnetization transfer spectroscopy. *Journal of Magnetic Resonance* **65**, 355-360.

- Behravan, G., Lycksell, P.-O., Larsson, G., 1997. Expression, purification and characterization of the homeodomain of rat ISL-1 protein. *Protein Engineering* **10**, 1327-1331.
- Benassayag, C., Seroude, L., Boube, M., Erard, M., Cribbs, D.L., 1997. A homeodomain point mutation of the *Drosophila* proboscipedia protein provokes eye loss independently of homeotic function. *Mechanisms of Development* **63**, 187-198.
- Berg, O.G., Winter, R.B., von Hippel, P.H., 1981. Diffusion-driven mechanisms of protein translocation on nucleic-acids: 1. Models and theory. *Biochemistry* **20**, 6929-6948.
- Billeter, M., Qian, Y.Q., Otting, G., Müller, M., Gehring, W.J., Wüthrich, K., 1993. Determination of the nuclear magnetic resonance solution structure of an *Antennapedia* homeodomain-DNA complex. *The Journal of Molecular Biology* **234**, 1084-1097.
- Bosshard, H.R., Marti, D.N., Jelesarov, I., 2004. Protein stabilization by salt bridges: concepts, experimental approaches and clarification of some misunderstandings. *Journal of Molecular Recognition* **17**, 1-16.
- Brickman, J.M., Clements, M., Tyrell, R., McNay, D., Woods, K., Warner, J., Stewart, A., Beddington, R.S.P., Dattani, M., 2001. Molecular effects of novel mutations in *Hesx1/HESX1* associated with human pituitary disorders. *Development* **128**, 5189-5199.
- Bullock, A.N., Henckel, J., Fersht, A.R., 2000. Quantitative analysis of residual folding and DNA binding in mutant p53 core domain: definition of mutant states for rescue in cancer therapy. *Oncogene* **19**, 1245-1256.
- Cañadillas, J.M.P., Tidow, H., Freund, S.M.V., Rutherford, T.J., Ang, H.C., Fersht, A.R., 2006. Solution structure of p53 core domain: Structural basis for its instability. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 2109-2114.
- Carra, J.H., Privalov, P.L., 1997. Energetics of folding and DNA binding of the MAT α 2 homeodomain. *Biochemistry* **36**, 526-535.
- Catron, K.M., Iler, N., Abate, C., 1993. Nucleotides flanking a conserved TAAT core dictate the DNA binding specificity of three murine homeodomain proteins. *Molecular and Cellular Biology* **13**, 2354-2365.
- Chakrabartty, A., Kortemme, T., Baldwin, R.L., 1994. Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. *Protein Science* **3**, 843-852.

- Clarke, N.D., Kissinger, C.R., Desjarlais, J., Gilliland, G.L., Pabo, C.O., 1994. Structural studies of the engrailed homeodomain. *Protein Science* **3**, 1779-1787.
- Clarke, N.D., 1995. Covariation of residues in the homeodomain sequence family. *Protein Science* **4**, 2269-2278.
- Cooper, A., Johnson, C.M., Lakey, J.H., Nöllmann, M., 2001. Heat does not come in different colours: entropy-enthalpy compensation, free energy windows, quantum confinement, pressure perturbation calorimetry, solvation and the multiple causes of heat capacity effects in biomolecular interactions. *Biophysical Chemistry* **93**, 215-230.
- Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., Kollman, P.A., 1995. A second generation force field for the simulation of proteins, nucleic acids and organic molecules. *The Journal of the American Chemical Society* **117**, 5179-5197.
- Cox, M., van Tilborg, P.J.A., de Laat, W., Boelens, R., van Leeuwen, H.C., van der Vliet, P.C., Kaptein, R., 1995. Solution structure of the Oct-1 POU homeodomain determined by NMR and restrained molecular dynamics. *Journal of Biomolecular NMR* **5**, 23-32.
- Dahiyat, B. I., Mayo, S.L., 1997. Probing the role of packing specificity in protein design. *Proceedings of the National Academy of Sciences USA* **94**, 10172-10177.
- Damante, G., Tell, G., Leonardi, A., Fogolari, F., Bortolotti, N., Di Lauro, R., Formisano, S., 1994. Analysis of the conformation and stability of rat TTF-1 homeodomain by circular dichroism. *FEBS Letters* **354**, 293-296.
- Dattani, M.T., Martinez-Barbera, J.P., Thomas, P.Q., Brickman, J.M., Gupta, R., Mårtensson, I.L., Toresson, H., Fox, M., Wales, J.K.H., Hindmarsh, P.C., Krauss, S., Beddington, R.S.P., Robinson, I.C.A.F., 1998. Mutations in the homeobox gene HESX1/Hesx1 associated with septo-optic dysplasia in human and mouse. *Nature Genetics* **19**, 125-133.
- Dave, V., Zhao, C., Yang, F., Tung, C.-S., Ma J., 2000. Reprogrammable recognition codes in bicoid homeodomain-DNA interaction. *Molecular and Cellular Biology* **20**, 7673-7684.
- Debye, P., Huckel, E., 1923. The interionic attraction theory of deviations from ideal behaviour in solution. *Physikalische Zeitschrift* **24**, 185-206.

- Delboni, L.F., Mande, S.C., Rentier-Delrue, F., Mainfroid, V., Turley, S., Vellieux, F.M.D., Martial, J.A., Hol, W.G.J., 1995. Crystal structure of recombinant triosephosphate isomerase from *Bacillus stearothermophilus*. An analysis of potential thermostability factors in six isomerases with known three-dimensional structures points to the importance of hydrophobic interactions. *Protein Science* **4**, 2594-2604.
- D'Elia, A.V., Tell, G., Paron, I., Pellizari, L., Lonigro, R., Damante, G., 2001. Missense mutations of human homeoboxes: a review. *Human Mutation* **18**, 361-374.
- Del Vecchio, P., Carullo, P., Barone, G., Pagano, B., Graziano, G., Iannetti, A., Acquaviva, R., Leonardi, A., Formisano, S., 2008. Conformational stability and DNA binding energetics of the rat thyroid transcription factor 1 homeodomain. *Proteins: Structure, Function and Genetics* **70**, 748-760.
- Dill, K.A., Bromberg, S., Yue, K.Z., Fiebig, K.M., Yee, D.P., Thomas, P.D., Chan, H.S., 1995. Principles of protein folding -- A perspective from simple exact models. *Protein Science* **4**, 561-602.
- Di Nardo, A.A., Larson, S.M., Davidson, A.R., 2003. The relationship between conservation, thermodynamic stability, and function in the SH3 domain hydrophobic core. *The Journal of Molecular Biology* **333**, 641-655.
- Dominy, B.N., Perl, D., Schmid, F.X., Brooks III, C.L., 2002. The effects of ionic strength on protein stability: the cold shock protein family. *The Journal of Molecular Biology* **319**, 541-554.
- Dragan, A.I., Li, Z., Makeyeva, E.N., Milgotina, E.I., Liu, Y., Crane-Robinson, C., Privalov, P.L., 2006. Forces driving the binding of homeodomains to DNA. *Biochemistry* **45**, 141-151.
- Dunitz, J.D., 1995. Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions. *Chemistry & Biology* **2**, 709-712.
- Ekker, S.C., von Kessler, D.P., Beachy, P.A., 1992. Differential DNA sequence recognition is a determinant of specificity in homeotic gene action. *The EMBO Journal* **11**, 4059-4072.
- Elcock, A.H., 1998. The stability of salt bridges at high temperatures: implications for hyperthermophilic proteins. *The Journal of Molecular Biology* **284**, 489-502.
- Esposito, G., Fogolari, F., Damante, G., Formisano, S., Tell, G., Leonardi, A., Di Lauro, R., Viglino, P., 1996. Analysis of the solution structure of the homeodomain of rat thyroid transcription factor 1 by ^1H -NMR

- spectroscopy and restrained molecular mechanics. *European Journal of Biochemistry* **241**, 101-113.
- Fabbro, D., Tell, G., Leonardi, A., Pellizzari, L., Pucillo, C., Lonigro, R., Formisano, S., Damante, G., 1996. In the TTF-1 homeodomain the contribution of several amino acids to DNA recognition depends on the bound sequence. *Nucleic Acids Research* **24**, 3283-3288.
- Fausti, S., Weiler, S., Cuniberti, C., Hwang, K.J., No, K.T., Gruschus, J.M., Perico, A., Nirenberg, M., Ferretti, J.A., 2001. Backbone dynamics for the wild type and a double H52R/T56W mutant of the vnd/NK-2 homeodomain from *Drosophila melanogaster*. *Biochemistry* **40**, 12004-12012.
- Fersht, A., 1999. *Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding*. W. H. Freeman and Company, New York.
- Fields, P.A., 2001. Review: Protein function at thermal extremes: balancing stability and flexibility. *Comparative Biochemistry and Physiology Part A* **129**, 417-431.
- Fodor, E., Mack, J.W., Maeng, J.S., Ju, J.H., Lee, H.S., Gruschus, J.M., Ferretti, J.A., Ginsburg, A., 2005. Cardiac-specific Nkx2.5 homeodomain: conformational stability and specific DNA binding of Nkx2.5(C56S). *Biochemistry* **44**, 12480-12490.
- Fraenkel, E., Pabo, C.O., 1998. Comparison of X-ray and NMR structures for the Antennapedia homeodomain-DNA complex. *Nature Structural Biology* **5**, 692-697.
- Fraenkel, E., Rould, M.A., Chambers, K.A., Pabo, C.O., 1998. Engrailed homeodomain-DNA complex at 2.2 Å resolution: a detailed view of the interface and comparison with other engrailed structures. *The Journal of Molecular Biology* **284**, 351-361.
- Frankel, A.D., Berg, J.M., Pabo, C.O., 1987. Metal-dependent folding of a single zinc finger from transcription factor IIIA. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 4841-4845.
- Frasch, M., Warrior, R., Tugwood, J., Levine, M., 1988. Molecular analysis of even-skipped mutants in *Drosophila* development. *Genes & Development* **2**, 1824-1838.
- Freire, E., 2001. The thermodynamic linkage between protein structure, stability, and function. *Methods in Molecular Biology* **168**, 37-68.
- Freund, C.L., Gregory-Evans, C.Y., Furukawa, T., Papaioannou, M., Looser, J., Ploder, L., Bellingham, J., Ng, D., Herbrick, J.-A.S., Duncan, A., Scherer,

- S.W., Tsui, L.-C., Loutradis-Anagnostou, A., Jacobson, S.G., Cepko, C.L., Bhattacharya, S.S. McInnes, R.R., 1997. Cone-rod dystrophy due to mutations in a novel photoreceptor-specific homeobox gene (*CRX*) essential for maintenance of the photoreceptor. *Cell* **91**, 543-553.
- Gage, P.J., Suh, H., Camper, S.A., 1999. Dosage requirement of *Pitx2* for development of multiple organs. *Development* **126**, 4643-4651.
- Gehring, W.J., Qian, Y.Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A.F., Resendez-Perez, D., Affolter, M., Otting, G., Wüthrich, K., 1994a. Homeodomain-DNA recognition. *Cell* **78**, 211-223.
- Gehring, W.J., Affolter, M., Burglin, T., 1994b. Homeodomain proteins. *Annual Review of Biochemistry* **63**, 487-526.
- Givaty, O., Levy, Y., 2009. Protein sliding along DNA: dynamics and structural characterization. *The Journal of Molecular Biology* **385**, 1087-1097.
- Gloss, L.M., Placek, B.J., 2002. The effect of salts on the stability of the H2A-H2B histone dimer. *Biochemistry* **41**, 14951-14959.
- Goldman, A., 1995. How to make my blood boil. *Structure* **3**, 1277-1279.
- Gonzalez, M., Weiler, S., Ferretti, J.A., Ginsburg, A., 2001. The vnd/NK-2 homeodomain: thermodynamics of reversible unfolding and DNA binding for wild-type and with residue replacements H52R and H52R/T56W in Helix III. *Biochemistry* **40**, 4923-4931.
- Grant, R.A., Rould, M.A., Klemm, J.D., Pabo, C.O., 2000. Exploring the role of glutamine 50 in the homeodomain-DNA interface: crystal structure of engrailed (Gln50→Ala) complex at 2.0 Å. *Biochemistry* **39**, 8187-8192.
- Grigeliuniene, G., Eklöf, O., Ivarsson, S.A., Westphal, O., Neumeyer, L., Kedra, D., Dumanski, J., Hagenäs, L., 2000. Mutations in short stature homeobox containing gene (*SHOX*) in dyschondrosteosis but not in hypochondroplasia. *Human Genetics* **107**, 145-149.
- Gruschus, J.M., Tsao, D.H.H., Wang, L.-H., Nirenberg, M., Ferretti, J.A., 1997. Interactions of the vnd/NK-2 homeodomain with DNA by nuclear magnetic resonance spectroscopy: basis of binding specificity. *Biochemistry* **36**, 5372-5380.
- Gruschus, J.M., Tsao, D.H.H., Wang, L.-H., Nirenberg, M., Ferretti, J.A., 1999. The three-dimensional structure of the vnd/NK-2 homeodomain-DNA complex by NMR spectroscopy. *The Journal of Molecular Biology* **289**, 529-545.

- Güntert, P., Qian, Y.Q., Otting, G., Müller, M., Gehring, W.J., Wüthrich, K., 1991. Structure determination of the *Antp*(C39→S) homeodomain from nuclear magnetic resonance data in solution using a novel strategy for the structure calculation with the programs DIANA, CALIBA, HABAS and GLOMSA. *The Journal of Molecular Biology* **217**, 531-540.
- Güntert, P., Mumenthaler, C., Wüthrich, K., 1997. Torsion angle dynamics for NMR structure calculation with the new program DYANA. *The Journal of Molecular Biology* **273**, 283-298.
- Hagan, D.M., Ross, A.J., Strachan, T., Lynch, S.A., Ruiz-Perez, V., Wang, Y.M., Scambler, P., Custard, E., Reardon, W., Hassan, S., Muenke, M., Nixon, P., Papapetrou, C., Winter, R.M., Edwards, Y., Morrison, K., Barrow, M., Cordier-Alex, M.P., Correia, P., Galvin-Parton, P.A., Gaskill, S., Gaskin, K.J., Garcia-Minaur, S., Gereige, R., Hayward, R., Homfray, T., McKeown, C., Murday, V., Plauchu, H., Shannon, N., Spitz, L., Lindsay, S., 2000. Mutation analysis and embryonic expression of the *HLXB9* Currarino syndrome gene. *The American Journal of Human Genetics* **66**, 1504-1515.
- Heberlein, U., Penton, A., Falsafi, S., Hackett, D., Rubin, G.M., 1994. The C-terminus of the homeodomain is required for functional specificity of the *Drosophila* rough gene. *Mechanisms of Development* **48**, 35-49.
- Hendsch, Z.S., Tidor, B., 1994. Do salt bridges stabilize proteins? A continuum electrostatic analysis. *Protein Science* **3**, 211-226.
- Higuchi, R., Krummel, B., Saiki, R.K., 1988. A general method of *in vitro* preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Research* **16**, 7351-7367.
- Hirsch, J.A., Aggarwal, A.K., 1995. Structure of the even-skipped homeodomain complexed to AT-rich DNA: new perspectives on homeodomain specificity. *The EMBO Journal* **14**, 6280-6291.
- Horovitz, A., Serrano, L., Avron, B., Bycroft, M., Fersht, A.R., 1990. Strength and co-operativity of contributions of surface salt bridges to protein stability. *The Journal of Molecular Biology* **216**, 1031-1044.
- Hovde, S., Abate-Shen, C., Geiger, J.H., 2001. Crystal structure of the *Msx-1* homeodomain/DNA complex. *Biochemistry* **40**, 12013-12021.
- Hu, G., Vastardis, H., Bendall A.J., Wang, Z., Logan, M., Zhang, H., Nelson, C., Stein, S., Greenfield, N., Seidman, C.E., Seidman, J.G., Abate-Shen, C., 1998. Haploinsufficiency of *MSX1*: a mechanism for selective tooth agenesis. *Molecular and Cellular Biology* **18**, 6044-6051.

- Ippel, H., Larsson, G., Behravan, G., Zdunek, J., Lundqvist, M., Schleucher, J., Lycksell, P.-O., Wijmenga, S., 1998. The solution structure of the homeodomain of the rat insulin-gene enhancer protein Isl-1. Comparison with other homeodomains. *The Journal of Molecular Biology* **288**, 689-703.
- Isaac, V.E., Sciavolino, P., Abate, C., 1995. Multiple amino acids determine the DNA binding specificity of the Msx-1 homeodomain. *Biochemistry* **34**, 7127-7134.
- Iurcu-Mustata, G., Van Belle, D., Wintjens, R., Prévost, M., Rooman, M., 2001. Role of salt bridges in homeodomains investigated by structural analyses and molecular dynamics simulations. *Biopolymers* **59**, 145-159.
- Iwahara, J., Clore, G.M., 2006. Detecting transient intermediates in macromolecular binding by paramagnetic NMR. *Nature* **440**, 1227-1230.
- Iwahara, J., Zweckstetter, M., Clore, G.M., 2006. NMR structural and kinetic characterization of a homeodomain diffusing and hopping on nonspecific DNA. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 15062-15067.
- Jabs, E.W., Müller, U., Li, X., Ma, L., Luo, W., Haworth, I.S., Klisak, I., Sparkes, R., Warman, M.L., Mulliken, J.B., Snead, M.L., Maxson, R., 1993. A mutation in the homeodomain of the human *MSX2* gene in a family affected with autosomal dominant craniosynostosis. *Cell* **75**, 443-450.
- Jackson, S.E., Moracci, M., elMasry, N., Johnson, C.M., Fersht, A.R., 1993. Effect of cavity-creating mutations in the hydrophobic core of chymotrypsin inhibitor 2. *Biochemistry* **32**, 11259-11269.
- Jemth, P., Gianni, S., 2007. PDZ domains: folding and binding. *Biochemistry* **46**, 8701-8708.
- Julenius, K., Thulin, E., Linse, S., Finn, B.E., 1998. Hydrophobic core substitutions in calbindin D9k: effects on stability and structure. *Biochemistry* **37**, 8915-8925.
- Kajander, T., Kahn, P.C., Passila, S.H., Cohen, D.C., Lethiö, L., Adolfsen, W., Warwicker, J., Schell, U., Goldman, A., 2000. Buried charged surface in proteins. *Structure* **8**, 1203-1214.
- Kao, Y.H., Fitch, C.A., Bhattacharya, S., Sarkisian, C.J., Lecomte, J.T.J., García-Moreno, B., 2000. Salt effects on ionization equilibria of histidines in myoglobin. *Biophysical Journal* **79**, 1637-1654.

- Karshikoff, A., Ladenstein, R., 2001. Ion pairs and the thermotolerance of proteins from hyperthermophiles: a 'traffic rule' for hot roads. *Trends in Biochemical Sciences* **26**, 550-556.
- Kazanskaya, O.V., Severtzova, E.A., Barth, K.A., Ermakova, G.V., Lukyanov, S.A., Benyumov, A.O., Pannese, M., Boncinelli, E., Wilson, S.W., Zaráisky, A.G., 1997. *Anf*: a novel class of vertebrate homeobox genes expressed at the anterior end of the main embryonic axis. *Gene* **200**, 25-34.
- Ke, A., Mathias, J.R., Vershon, A.K., Wolberger, C., 2002. Structural and thermodynamic characterization of the DNA binding properties of a triple alanine mutant of MAT α 2. *Structure* **10**, 961-971.
- Kissinger, C.R., Liu, B., Martín-Blanco, E., Kornberg, T.B., Pabo, C.O., 1990. Crystal structure of an engrailed homeodomain-DNA complex at 2.8 Å resolution: a framework for understanding homeodomain-DNA interactions. *Cell* **63**, 579-590.
- Kragelund, B.B., Jonsson, M., Bifulco, G., Chazin, W.J., Nilsson, H., Finn, B.E., Linse, S., 1998. Hydrophobic core substitutions in calbindin D9k: effects on Ca²⁺ binding and dissociation. *Biochemistry* **37**, 8926-8937.
- Kulak, S.C., Kozłowski, K., Semina, E.V., Pearce, W.G., Walter, M.A., 1998. Mutation in the *RIEG1* gene in patients with iridogoniodysgenesis syndrome. *Human Molecular Genetics* **7**, 1113-1117.
- Kumar, A., Ernst, R.R., Wüthrich, K., 1980. A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochemical and Biophysical Research Communications* **95**, 1-6.
- Kumar, S., Nussinov, R., 1999. Salt bridge stability in monomeric proteins. *The Journal of Molecular Biology* **293**, 1241-1255.
- Kumar, S., Nussinov, R., 2001. How do thermophilic proteins deal with heat? *Cellular and Molecular Life Sciences* **58**, 1216-1233.
- Kumar, S., Tsai, C.J., Nussinov, R., 2003. Temperature range of thermodynamic stability for the native state of reversible two-state proteins. *Biochemistry* **42**, 4864-4873.
- Laughon, A., 1991. DNA binding specificity of homeodomains. *Biochemistry* **30**, 11357-11367.
- Lee, K.K., Fitch, C.A., García-Moreno, B., 2002a. Distance dependence and salt sensitivity of pairwise, coulombic interactions in a protein. *Protein Science* **11**, 1004-1016.

- Lee, K.K., Fitch, C.A., Lecomte, J.T.J., García-Moreno, B., 2002b. Electrostatic effects in highly charged proteins: salt sensitivity of pK_a values of histidines in staphylococcal nuclease. *Biochemistry* **41**, 5656-5667.
- Li, T., Stark, M.R., Johnson, A.D., Wolberger, C., 1995. Crystal structure of the MAT α 1/MAT α 2 homeodomain heterodimer bound to DNA. *Science* **270**, 262-269.
- Lim, W.A., Sauer, R.T., 1991. The role of internal packing interactions in determining the structure and stability of a protein. *The Journal of Molecular Biology* **219**, 359-376.
- Lim, W.A., Farruggio, D.C., Sauer, R.T., 1992. Structural and energetic consequences of disruptive mutations in a protein core. *Biochemistry* **31**, 4324-4333.
- Lindberg, M.J., Bystrom, R., Boknas, N., Andersen, P.M., Oliveberg, M., 2005. Systematically perturbed folding patterns of amyotrophic lateral sclerosis (ALS)-associated SOD1 mutants. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 9754-9759.
- Lindman, S., Xue, W.-F., Szczepankiewicz, O., Bauer, M.C., Nilsson, H., Linse, S., 2006. Salting the charged surface: pH and salt dependence of protein G B1 stability. *Biophysical Journal* **90**, 2911-2921.
- Longo, A., Guanga, G.P., Rose, R.B., 2007. Structural basis for induced fit mechanisms in DNA recognition by the Pdx1 homeodomain. *Biochemistry* **46**, 2948-2957.
- Lyu, P.C., Gans, P.J., Kallenbach, N.R., 1992. Energetic contribution of solvent-exposed ion pairs to alpha-helix structure. *The Journal of Molecular Biology* **223**, 343-350.
- Ma, L., Golden, S., Wu, L., Maxson, R., 1996. The molecular basis of Boston-type craniosynostosis: the Pro148→His mutation in the N-terminal arm of the MSX2 homeodomain stabilizes DNA binding without altering nucleotide sequence preferences. *Human Molecular Genetics* **5**, 1915-1920.
- Maerkl, S.J., Quake, S.R., 2007. A systems approach to measuring the binding energy landscapes of transcription factors. *Science* **315**, 233-237.
- Makhatadze, G.I., Loladze, V.V., Ermolenko, D.N., Chen, X., Thomas, S.T., 2003. Contribution of surface salt bridges to protein stability: guidelines for protein engineering. *The Journal of Molecular Biology* **327**, 1135-1148.
- Marion, D., Wüthrich, K., 1983. Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of ^1H - ^1H

- spin-spin coupling constants in proteins. *Biochemical and Biophysical Research Communications* **113**, 967-974.
- Marqusee, S., Baldwin, R.L., 1987. Helix stabilization by Glu⁻...Lys⁺ salt bridges in short peptides of *de novo* design. *Proceedings of the National Academy of Sciences USA* **84**, 8898-8902.
- Marqusee, S., Sauer, R.T., 1994. Contribution of a hydrogen bond / salt bridge network to the stability of secondary and tertiary structures in lambda repressor. *Protein Science* **3**, 2217-2225.
- Marshall, S.A., Mayo, S.L., 2001. Achieving stability and conformational specificity in designed proteins *via* binary patterning. *The Journal of Molecular Biology* **305**, 619-631.
- Marshall, S.A., Morgan, C.S., Mayo, S.L., 2002. Electrostatics significantly affect the stability of designed homeodomain variants. *The Journal of Molecular Biology* **316**, 189-199.
- Marti, D.N., Jelesarov, I., Bosshard, H.R., 2000. Interhelical ion pairing in coiled coils: solution structure of a heterodimeric leucine zipper and determination of pK_a values of Glu side chains. *Biochemistry* **39**, 12804-12818.
- Marti, D.N., Bosshard, H.R., 2003. Electrostatic interactions in leucine zippers: thermodynamic analysis of the contributions of Glu and His residues and the effect of mutating salt bridges. *The Journal of Molecular Biology* **330**, 621-637.
- Martin, A., Kather, I., Schmid, F.X., 2002. Origins of the high stability of an *in vitro*-selected cold-shock protein. *The Journal of Molecular Biology* **318**, 1341-1349.
- Mason, J.M., Schmitz, M.A., Müller, K.M., Arndt, K.M., 2006. Semirational design of Jun-Fos coiled coils with increased affinity: Universal implications for leucine zipper prediction and design. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 8989-8994.
- de la Mata, I., García, J.L., González, C., Menéndez, M., Cañada, J., Jiménez-Barbero, J., Asensio, J.L., 2002. The impact of R53C mutation on the three-dimensional structure, stability and DNA-binding properties of the Human Hesx-1 homeodomain. *ChemBioChem* **3**, 726-740.
- Mathews, C.K., van Holde, K.E., Ahern, K.G., 2000. *Biochemistry* (3ª edición). Prentice Hall, Upper Saddle River, NJ.
- Max, K.E.A., Wunderlich, M., Roske, Y., Schmid, F.X., Heinemann, U., 2007. Optimized variants of the cold shock protein from *in vitro* selection:

- structural basis of their high thermostability. *The Journal of Molecular Biology* **369**, 1087-1097.
- Mayer, S., Rüdiger, S., Ang, H.C., Joerger, A.C., Fersht, A.R., 2007. Correlation of levels of folded recombinant p53 in *Escherichia coli* with thermodynamic stability *in vitro*. *The Journal of Molecular Biology* **372**, 268-276.
- Mayor, U., Johnson, C.M., Daggett, V., Fersht, A.R., 2000. Protein folding and unfolding in microseconds to nanoseconds by experiment and simulation. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 13518-13522.
- Mayor, U., Guydosh, N.R., Johnson, C.M., Grossmann, J.G., Sato, S., Jas, G.S., Freund, S.M.V., Alonso, D.O.V., Daggett, V., Fersht, A.R., 2003. The complete folding pathway of a protein from nanoseconds to microseconds. *Nature* **421**, 863-867.
- Mayr, L.M., Landt, O., Hahn, U., Schmid, F.X., 1993. Stability and folding kinetics of Ribonuclease T₁ are strongly altered by the replacement of *Cis*-proline 39 with alanine. *The Journal of Molecular Biology* **231**, 897-912.
- Meier-Andrejszki, L., Bjelić, S., Naud, J.-F., Lavigne, P., Jelesarov, I., 2007. Thermodynamic of b-HLH-LZ protein binding to DNA: The energetic importance of protein-DNA contacts in site-specific E-Box recognition by the complete gene product of the Max p21 transcription factor. *Biochemistry* **46**, 12427-12440.
- Morita, E.H., Shirakawa, M., Hayashi, F., Imagawa, M., Kyogoku, Y., 1993. Secondary structure of the *oct*-3 POU homeodomain as determined by ¹H-¹⁵N NMR spectroscopy. *FEBS Letters* **321**, 107-110.
- Mueller, U., Perl, D., Schmid, F.X., Heinemann, U., 2000. Thermal stability and atomic-resolution crystal structure of the *Bacillus caldolyticus* cold shock protein. *The Journal of Molecular Biology* **297**, 975-988.
- Müller, M., Affolter, M., Leupin, W., Otting, G., Wüthrich, K., Gehring, W.J., 1988. Isolation and sequence-specific DNA binding of the *Antennapedia* homeodomain. *The EMBO Journal* **7**, 4299-4304.
- Myers, J.K., Pace, C.N., Scoltz, J.M., 1995. Denaturant *m* values and heat capacity changes: relation to changes in accesible surface areas of protein unfolding. *Protein Science* **4**, 2138-2148.
- Nanda, V., Brand, L., 2000. Aromatic interactions in homeodomains contribute to the low quantum yield of a conserved, buried tryptophan. *Proteins: Structure, Function and Genetics* **40**, 112-125.

- Nozaki, Y., Tanford, C., 1967. Examination of titration behavior. *Methods in Enzymology* **11**, 715-734.
- Ó Fágáin, C., 1995. Understanding and increasing protein stability. *Biochimica et Biophysica Acta* **1252**, 1-14.
- Ohlendorf, D.H., Anderson, W.F., Fisher, R.G., Takeda, Y., Matthews, B.W., 1982. The molecular basis of DNA-protein recognition inferred from the structure of cro repressor. *Nature* **298**, 718-723.
- Olufsen, M., Papaleo, E., Smalås, A.O., Brandsdal, B.O., 2007. Ion pairs and their role in modulating stability of cold- and warm-active uracil DNA glycosylase. *Proteins: Structure, Function and Genetics* **71**, 1219-1230.
- O'Neil, K.T., DeGrado, W.F., 1990. A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science* **250**, 646-651.
- Otting, G., Qian, Y.Q., Muller, M., Affolter, M., Gehring, W.J., Wüthrich, K., 1988. Secondary structure determination for the Antennapedia homeo-domain by nuclear magnetic resonance and evidence for a helix-turn-helix motif. *The EMBO Journal* **7**, 4305-4309.
- Otting, G., Qian, Y.Q., Billeter, M., Müller, M., Affolter, M., Gehring, W.J., Wüthrich, K., 1990. Protein-DNA contacts in the structure of a homeodomain-DNA complex determined by nuclear magnetic resonance spectroscopy in solution. *The EMBO Journal* **9**, 3085-3092.
- Pace, C.N., 1986. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods in Enzymology* **131**, 266-280.
- Pace, C.N., Scholtz, J.M., 1998. A helix propensity scale based on experimental studies of peptides and proteins. *Biophysical Journal* **75**, 422-427.
- Pace, C.N., Alston, R.W., Shaw, K.L., 2000. Charge-charge interactions influence the denatured state ensemble and contribute to protein stability. *Protein Science* **9**, 1395-1398.
- Pack, S.P., Yoo, Y.J., 2003. Protein thermostability: structure-based difference of residual properties between thermophilic and mesophilic proteins. *Journal of Molecular Catalysis B: Enzymatic* **26**, 257-264.
- Pack, S.P., Yoo, Y.J., 2004. Protein thermostability: structure-based difference of amino acid between thermophilic and mesophilic proteins. *Journal of Biotechnology* **111**, 269-277.

- Padmanabhan, S., Jiménez, M.A., González, C., Sanz, J.M., Giménez-Gallego, G., Rico, M., 1997. Three-dimensional solution structure and stability of phage 434 Cro protein. *Biochemistry* **36**, 6424-6436.
- Padmanabhan, S., Laurents, D.V., Fernández, A.M., Elias-Arnanz, M., Ruiz-Sanz, J., Mateo, P.L., Rico, M., Filimonov, V.V., 1999. Thermodynamic analysis of the structural stability of phage 434 Cro protein. *Biochemistry* **38**, 15536-15547.
- Percival-Smith, A., Müller, M., Affolter, M., Gehring, W.J., 1990. The interaction with DNA of wild-type and mutant *fushi tarazu* homeodomains. *The EMBO Journal* **9**, 3967-3974.
- Perl, D., Mueller, U., Heinemann, U., Schmid, F.X., 2000. Two exposed amino acid residues confer thermostability on a cold shock protein. *Nature Structural Biology* **7**, 380-383.
- Pernasetti, F., Milner, R.D.G., Al Ashwal, A.A.Z., de Zegher, F., Chavez, V.M., Muller, M., Martial, J.A., 1998. Pro239Ser: a novel recessive mutation of the Pit-1 gene in seven Middle Eastern children with growth hormone, prolactin, and thyrotropin deficiency. *Journal of Clinical Endocrinology and Metabolism* **83**, 2079-2083.
- Perutz, M.F., Raidt, H., 1975. Stereochemical basis of heat stability in bacterial ferredoxines and in haemoglobin A2. *Nature* **255**, 256-259.
- Perveen, R., Lloyd, I.C., Clayton-Smith, J., Churchill, A., van Heyningen, V., Hanson, I., Taylor, D., McKeown, C., Super, M., Kerr, B., Winter, R., Black, G.C.M., 2000. Phenotypic variability and asymmetry of Rieger syndrome associated with *PITX2* mutations. *Investigative Ophthalmology & Visual Science* **41**, 2456-2460.
- Pervushin, K., Billeter, M., Siegal, G., Wüthrich, K., 1996. Structural role of a buried salt bridge in the 434 repressor DNA-binding domain. *The Journal of Molecular Biology* **264**, 1002-1012.
- Phillips, C.L., Vershon, A.K., Johnson, A.D., Dahlquist, F.W., 1991. Secondary structure of the homeo domain of yeast $\alpha 2$ repressor determined by NMR spectroscopy. *Genes & Development* **5**, 764-772.
- Piotto, M., Saudek, V., Sklenar, V., 1992. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *Journal of Biomolecular NMR* **2**, 661-665.
- Privalov, P.L., Jelesarov, I., Read, C.M., Dragan, A.I., Crane-Robinson, C., 1999. The energetics of HMG box interactions with DNA: thermodynamics of

- the DNA binding of the HMG box from mouse Sox-5. *The Journal of Molecular Biology* **294**, 997-1013.
- Privalov, P.L., Dragan, A.I., Crane-Robinson, C., Breslauer, K.J., Remeta, D.P., Minetti, C.A., 2007. What drives proteins into the major or minor grooves of DNA? *The Journal of Molecular Biology* **365**, 1-9.
- Qian, Y.Q., Furukubo-Tokunaga, K., Resendez-Perez, D., Müller, M., Gehring, W.J., Wüthrich, K., 1994a. Nuclear magnetic resonance solution structure of the fushi tarazu homeodomain from *Drosophila* and comparison with the Antennapedia homeodomain. *The Journal of Molecular Biology* **238**, 333-345.
- Qian, Y.Q., Resendez-Perez, D., Gehring, W.J., Wüthrich, K., 1994b. The des(1-6)Antennapedia homeodomain: Comparison of the NMR solution structure and the DNA-binding affinity with the intact Antennapedia homeodomain. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 4091-4095.
- Radovick, S., Nations, M., Du, Y., Berg, L.A., Weintraub, B.D., Wondisford, F.E., 1992. A mutation in the POU-homeodomain of Pit-1 responsible for combined pituitary hormone deficiency. *Science* **257**, 1115-1118.
- Redfield, A.G., Kuntz, S.D., 1975. Quadrature Fourier NMR detection: simple multiplex for dual detection and discussion. *Journal of Magnetic Resonance* **19**, 250-254.
- de los Rios, M.A., Plaxco, K.W., 2005. Apparent Debye-Huckel electrostatic effects in the folding of a simple, single domain protein. *Biochemistry* **44**, 1243-1250.
- Sakai-Kato, K., Umezawa, Y., Mikoshiba, K., Aruga, J., Utsunomiya-Tate, N., 2009. Stability of folding structure of Zic zinc finger proteins. *Biochemical and Biophysical Research Communications* **384**, 362-365.
- Sali, D., Bycroft, M., Fersht, A.R., 1991. Surface electrostatic interactions contribute little to stability of Barnase. *The Journal of Molecular Biology* **220**, 779-788.
- Santoro M.M., Bolen, D.W., 1988. Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of Phenylmethanesulfonyl α -Chymotrypsin using different denaturants. *Biochemistry* **27**, 8063-8068.
- Sato, K., Simon, M.D., Levin, A.M., Shokat, K.M., Weiss, G.A., 2004. Dissecting the Engrailed homeodomain-DNA interaction by phage-displayed shotgun scanning. *Chemistry & Biology* **11**, 1017-1023.

- Scheek, R.M., Vangunsteren, W.F., Kaptein, R., 1989. Molecular dynamics simulation techniques for determination of molecular structures from nuclear magnetic resonance data. *Methods in Enzymology* **177**, 204-218.
- Scheffer, I.E., Wallace, R.H., Phillips, F.L., Hewson, P., Reardon, K., Parasivam, G., Stromme, P., Berkovic, S.F., Gecz, J., Mulley, J.C., 2002. X-linked myoclonic epilepsy with spasticity and intellectual disability. *Neurology* **59**, 348-356.
- Seimiya, M., Kurosawa, Y., 1996. Kinetics of binding of Antp homeodomain to DNA analyzed by measurements of surface plasmon resonance. *FEBS Letters* **398**, 279-284.
- Semina, E.V., Reiter, R., Leysens, N.J., Alward, W.L.M., Small, K.W., Datson, N.A., Siegel-Bartelt, J., Bierke-Nelson, D., Bitoun, P., Zabel, B.U., Carey, J.C., Murray, J.C., 1996. Cloning and characterization of a novel *bicoid*-related homeobox transcription factor gene, *RIEG*, involved in Rieger syndrome. *Nature Genetics* **14**, 392-399.
- Serrano, L., Horovitz, A., Avron, B., Bycroft, M., Fersht, A.R., 1990. Estimating the contribution of engineered surface electrostatic interactions to protein stability by using double-mutant cycles. *Biochemistry* **29**, 9343-9352.
- Serrano, L., Kellis, J.T. Jr, Cann, P., Matouschek, A., Fersht, A.R., 1992. The folding of an enzyme. II. Substructure of barnase and the contribution of different interactions to protein stability. *The Journal of Molecular Biology* **224**, 783-804.
- Shah, P.S., Hom, G.K., Ross, S.A., Lassila, J.K., Crowhurst, K.A., Mayo, S.L., 2007. Full-sequence computational design and solution structure of a thermostable protein variant. *The Journal of Molecular Biology* **372**, 1-6.
- Shang, Z., Isaac, V.E., Li, H., Patel, L., Catron, K.M., Curran, T., Montelione, G.T., Abate, C., 1994. Design of a "minimAl" homeodomain: The N-terminal arm modulates DNA binding affinity and stabilizes homeodomain structure. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 8373-8377.
- Shortle, D., Stites, W.E., Meeker, A.K., 1990. Contribution of the large hydrophobic amino acids to the stability of staphylococcal nuclease. *Biochemistry* **29**, 8033-8041.
- Sivaraja, M., Botfield, M.C., Mueller, M., Jancso, A., Weiss, M.A., 1994. Solution structure of a POU-specific homeodomain: 3D-NMR studies of human B-Cell transcription factor Oct-2. *Biochemistry* **33**, 9845-9855.

- Spek, E.J., Bui, A.H., Lu, M., Kallenbach, N.R., 1998. Surface salt bridges stabilize the GCN4 leucine zipper. *Protein Science* **7**, 2431-2437.
- Spolar, R.S., Record, M.T., 1994. Coupling of local folding to site-specific binding of proteins to DNA. *Science* **263**, 777-784.
- States, D., Haberkorn, R.A., Ruben, D.J., 1982. A two-dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants. *Journal of Magnetic Resonance* **48**, 286-292.
- Stollar, E.J., Mayor, U., Lovell, S.C., Federici, L., Freund, S.M.V., Fersht, A.R., Luisi, B.F., 2003. Crystal structures of Engrailed homeodomain mutants. *The Journal of Biological Chemistry* **278**, 43699-43708.
- Stouffer, A.L., Nanda, V., Lear, J.D., DeGrado, W.F., 2005. Sequence determinants of a transmembrane proton channel: an inverse relationship between stability and function. *The Journal of Molecular Biology* **347**, 169-179.
- Strømme, P., Mangelsdorf, M.E., Shaw, M.A., Lower, K.M., Lewis, S.M., Bruyere, H., Lütcherath, V., Gedeon, A.K., Wallace, R.H., Scheffer, I.E., Turner, G., Partington, M., Frints, S.G., Fryns, J.P., Sutherland, G.R., Mulley, J.C., Géczi, J., 2002. Mutations in the human ortholog of Aristaless cause X-linked mental retardation and epilepsy. *Nature Genetics* **30**, 441-445.
- Strop, P., Mayo, S.L., 2000. Contribution of surface salt bridges to protein stability. *Biochemistry* **39**, 1251-1255.
- Subramaniam, V., Jovin, T.M., Rivera-Pomar, R.V., 2001. Aromatic amino acids are critical for the stability of the Bicoid homeodomain. *The Journal of Biological Chemistry* **276**, 21506-21511.
- Swaroop, A., Wang, Q.-L., Wu, W., Cook, J., Coats, C., Xu, S., Chen, S., Zack, D.J., Sieving, P.A., 1999. Leber congenital amaurosis caused by a homozygous mutation (R90W) in the homeodomain of the retinal transcription factor CRX: direct evidence for the involvement of CRX in the development of photoreceptor function. *Human Molecular Genetics* **8**, 299-305.
- Takano, K., Ogasahara, K., Kaneda, H., Yamagata, Y., Fujii, S., Kanaya, E., Kikuchi, M., Oobatake, M., Yutani, K., 1995. Contribution of hydrophobic residues to the stability of human lysozyme: calorimetric studies and X-ray structural analysis of the five isoleucine to valine mutants. *The Journal of Molecular Biology* **254**, 62-76.

- Takano, K., Tsuchimori, K., Yamagata, Y., Yutani, K., 2000. Contribution of salt bridges near the surface of a protein to the conformational stability. *Biochemistry* **39**, 12375-12381.
- Teilum, K., Olsen, J.G., Kragelund, B.B., 2011. Protein stability, flexibility and function. *Biochimica et Biophysica Acta* **1814**, 969-976.
- Tell, G., Acquaviva, R., Formisano, S., Fogolari, F., Pucillo, C., Damante, G., 1999. Comparative stability analysis of the thyroid transcription factor and Antennapedia homeodomains: evidence for residue 54 controlling the structural stability of the recognition helix. *The International Journal of Biochemistry & Cell Biology* **31**, 1339-1353.
- Thomas, P.Q., Rathjen, P.D., 1992. HES-1, a novel homeobox gene expressed by murine embryonic stem cells, identifies a new class of homeobox genes. *Nucleic Acids Research* **20**, 5840.
- Thomas, P.Q., Johnson, B.V., Rathjen, J., Rathjen, P.D., 1995. Sequence, genomic organization, and expression of the novel Homeobox gene *Hesx-1*. *The Journal of Biological Chemistry* **270**, 3869-3875.
- Thomas, P.Q., Dattani, M.T., Brickman, J.M., McNay, D., Warne, G., Zacharin, M., Cameron, F., Hurst, J., Woods, K., Dunger, D., Stanhope, R., Forrest, S., Robinson, I.C.A.F., Beddington, R.S.P., 2001. Heterozygous *HESX1* mutations associated with isolated congenital pituitary hypoplasia and septo-optic dysplasia. *Human Molecular Genetics* **10**, 39-45.
- Travaglini-Allocatelli, C., Ivarsson, Y., Jemth, P., Gianni, S., 2009. Folding and stability of globular proteins and implications for function. *Current Opinion in Structural Biology* **19**, 3-7.
- Treisman, J., Gönczy, P., Vashishtha, M., Harris, E., Desplan, C., 1989. A single amino acid can determine the DNA binding specificity of homeodomain proteins. *Cell* **59**, 553-562.
- Tsao, D.H.H., Gruschus, J.M., Wang, L.H., Nirenberg, M., Ferretti, J.A., 1994. Elongation of helix III of the NK-2 homeodomain upon binding to DNA: a secondary structure study by NMR. *Biochemistry* **33**, 15053-15060.
- Tsao, D.H.H., Gruschus, J.M., Wang, L.H., Nirenberg, M., Ferretti, J.A., 1995. The three-dimensional solution structure of the NK-2 homeodomain from *Drosophila*. *The Journal of Molecular Biology* **251**, 297-307.
- Turner, E.C., Cureton, C.H., Weston, C.J., Smart, O.S., Allemann, R.K., 2004. Controlling the DNA binding specificity of bHLH proteins through intramolecular interactions. *Chemistry & Biology* **11**, 69-77.

- van Oostveen, J., Bijl, J., Raaphorst, F., Walboomers, J., Meijer, C., 1999. The role of homeobox genes in normal hematopoiesis and hematological malignancies. *Leukemia* **13**, 1675-1690.
- Vastardis, H., Karimbux, N., Guthua, S.W., Seidman, J.G., Seidman, C.E., 1996. A human *MSX1* homeodomain missense mutation causes selective tooth agenesis. *Nature Genetics* **13**, 417-421.
- Vuzman, D., Azia, A., Levy, Y., 2010. Searching DNA via a "Monkey Bar" mechanism: the significance of disordered tails. *The Journal of Molecular Biology* **396**, 674-684.
- Waldburger, C.D., Schildbach, J.F., Sauer, R.T., 1995. Are buried salt bridges important for protein stability and conformational specificity? *Nature Structural Biology* **2**, 122-128.
- Wallis, D.E., Roessler, E., Hehr, U., Nanni, L., Wiltshire, T., Richieri-Costa, A., Gillessen-Kaesbach, G., Zackai, E.H., Rommens, J., Muenke, M., 1999. Mutations in the homeodomain of the human *SIX3* gene cause holoprosencephaly. *Nature Genetics* **22**, 196-198.
- Weiler, S., Gruschus, J.M., Tsao, D.H.H., Yu, L., Wang, L.H., Nirenberg, M., Ferretti, J.A., 1998. Site-directed mutations in the vnd/NK-2 homeodomain. *The Journal of Biological Chemistry* **273**, 10994-11000.
- Wilson, D.S., Sheng, G., Jun, S., Desplan, C., 1996. Conservation and diversification in homeodomain-DNA interactions: A comparative genetic analysis. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 6886-6891.
- Wimley, W.C., Gawrisch, K., Creamer, T.P., White, S.H., 1996. Direct measurement of salt-bridge solvation energies using peptide model system: implications for protein stability. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 2985-2990.
- Wolberger, C., Vershon, A.K., Liu, B., Johnson, A.D., Pabo, C.O., 1991. Crystal structure of a MAT α 2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. *Cell* **67**, 517-528.
- Wolberger, C., 1996. Homeodomain interactions. *Current Opinion in Structural Biology* **6**, 62-68.
- Wu, W., Cogan, J.D., Pfäffle, R.W., Dasen, J.S., Frisch, H., O'Connell, S.M., Flynn, S.E., Brown, M.R., Mullis, P.E., Parks, J.S., Phillips III, J.A., Rosenfeld, M.G., 1998. Mutations in *PROP1* cause familial combined pituitary hormone deficiency. *Nature Genetics* **18**, 147-149.

- Wunderlich, M., Martin, A., Schmid, F.X., 2005. Stabilization of the cold shock protein CspB from *Bacillus subtilis* by evolutionary optimization of coulombic interactions. *The Journal of Molecular Biology* **347**, 1063-1076.
- Wüthrich, K., 1986. *NMR of proteins and nucleic acids*. John Wiley & Sons, Inc., New York.
- Wyman, J., 1948. Heme proteins. *Advances in Protein Chemistry* **4**, 407-531.
- Xiao, L., Honig, B., 1999. Electrostatic contributions to the stability of hyperthermophilic proteins. *The Journal of Molecular Biology* **289**, 1435-1444.
- Yang, A.S., Honig, B., 1993. On the pH dependence of protein stability. *The Journal of Molecular Biology* **231**, 459-474.
- Zollars, E.S., Marshall, S.A., Mayo, S.L., 2006. Simple electrostatic model improves designed protein sequences. *Protein Science* **15**, 2014-2018.

LISTA DE ABREVIATURAS

1D-RMN: resonancia magnética nuclear monodimensional	DQF-COSY: espectroscopía de correlación filtrada por doble cuanto (<i>double- quantum filtered correlation spectroscopy</i>)
1D: monodimensional	DSC: calorimetría diferencial de barrido (<i>differential scanning calorimetry</i>)
2D: bidimensional	DTT: DL-ditiotreitol
3D: tridimensional	DYANA: <i>dynamics algorithm for NMR applications</i> (algoritmo de dinámica para aplicaciones de RMN)
A: alanina	E: ácido glutámico
Å: angstroms	E: error
abs _{280 nm} : absorbancia a 280 nanómetros	EMSA: experimento de retardo en gel (<i>electrophoretic mobility shift assay</i>)
ADN: ácido desoxirribonucleico	Eng: Engrailed
Ala: alanina	F: fenilalanina
AMBER: <i>assisted model building with energy refinement</i> (construcción asistida de modelos con refinamiento energético)	F ₁ : eje de frecuencias 1 (dimensión indirecta)
Antp: Antennapedia	F ₂ : eje de frecuencias 2 (dimensión directa)
Arg: arginina	FPLC: cromatografía líquida de proteínas rápida (<i>fast protein liquid chromatography</i>)
ARX: homeobox relacionado con <i>aristaless</i> (<i>aristaless related homeobox</i>)	g: aceleración de la gravedad
Asn: asparagina	G: glicina
Asp: ácido aspártico	GdmCl: cloruro de guanidinio
C: cisteína	Giro II-III: giro que conecta las hélices II y III del homeodominio
cad.: cadena	Gln: glutamina
CBM: Centro de Biología Molecular	Glu: ácido glutámico
cDNA: ADN complementario	Gly: glicina
CIB: Centro de Investigaciones Biológicas	H: histidina
CNIO: Centro Nacional de Investigaciones Oncológicas	HD: homeodominio
COSY: espectroscopía de correlación (<i>correlation spectroscopy</i>)	HEPES: ácido 4-(2-hidroxietil)-1- -piperazinaetanosulfónico
CRX: homeobox de bastones y conos (<i>cone-rod homeobox</i>)	Hesx-1/HESX-1: homeobox expresado en células madre embrionarias 1 (<i>homeobox expressed in embryonic stem cells 1</i>)
CSIC: Consejo Superior de Investigaciones Científicas	His: histidina
Cys: cisteína	HLXB9: homeobox HB9
D: ácido aspártico	HOXA9: homeobox A9
DNA: <i>Deoxyribonucleic acid</i>	
dNTPs: desoxirribonucleótidos trifosfato	
DO _{600 nm} : densidad óptica a 600 nm	

HOXD9: homeobox D9
 I: isoleucina
 Ile: isoleucina
 IPTG: isopropil- β -D-tiogalactopiranosido
 ITC: *isothermal titration calorimetry*
 K: lisina
 K_b : constante de afinidad de unión (*binding*)
 L: leucina
 Lazo I-II: lazo que conecta las hélices I y II del homeodominio
 LB: caldo de lisogenia (*lysogeny broth*)
 Leu: leucina
 Lys: lisina
 M: metionina
 MALDI-TOF: espectrómetro de masas de tiempo de vuelo por desorción/ionización láser asistida por matriz (*matrix-assisted laser desorption/ionization time-of-flight mass spectrometer*)
 MAT α 2: proteína de apareamiento α 2 (*mating-type protein α 2*)
 mAU: miliunidades de absorbancia
 Met: metionina
 min: minutos
 MOPS: ácido 3-(N-morfolino)propanosulfónico
 MSX-1: msh homeobox 1
 MWCO: corte de peso molecular (*molecular weight cut off*)
 N: asparagina
 NK-2: antígeno asociado con las células *natural killer 2* (*natural killer cell-associated antigen 2*)
 NKX2.5: factor de transcripción relacionado con NK-2, locus 5
 NMR: *nuclear magnetic resonance*
 NOE: efecto nuclear Overhauser (*nuclear Overhauser effect*)
 NOESY: espectroscopía de efecto nuclear Overhauser (*nuclear Overhauser effect spectroscopy*)
 Oct-2: Proteína de unión a octámeros 2 (*Octamer-binding protein 2*)
 P: prolina
 p/v: peso/volumen
 pb: pares de bases
 PCR: reacción en cadena de la polimerasa (*polymerase chain reaction*)
 PDB: Protein Data Bank
 PDX-1: homeobox pancreático y duodenal 1 (*pancreatic and duodenal homeobox 1*)
 PEI: polietilenimina
 Phe: fenilalanina
 PITX2: homeobox de la pituitaria 2 (*pituitary homeobox 2*)
 pK_a^F : pK_a en la proteína plegada (*folded*)
 pK_a^U : pK_a en la proteína desplegada (*unfolded*)
 PMSF: fluoruro de fenilmetilsulfonilo
 ppm: partes por millón
 Pro: prolina
 PROP1: proteína homeobox profeta de Pit-1 (*homeobox protein prophet of Pit-1*)
 psi: libras por pulgada cuadrada
 Pt: precipitado
 Q: glutamina
 R: arginina
 Refs.: referencias
 REM: minimización restringida de energía (*restrained energy minimization*)
 RMD: dinámica molecular restringida (*restrained molecular dynamics*)
 RMN: resonancia magnética nuclear
 RMSD: desviación cuadrática media (*root-mean-square deviation*)
 rpm: revoluciones por minuto
 S: serina
 SDS-PAGE: electroforesis en gel de poliacrilamida con dodecilsulfato sódico (*sodium dodecyl sulfate polyacrylamide gel electrophoresis*)
 Ser: serina
 SHOX: homeobox de estatura baja (*short stature homeobox*)
 SIX3: homólogo 3 del homeobox *sine oculis* (*sine oculis homeobox homolog 3*)
 SN: sobrenadante
 SOD1: dismutasa de superóxidos 1 (*superoxide dismutase 1*)
 SPR: resonancia de plasmones superficiales (*surface plasmon resonance*)
 T: treonina
 T_{amb} : temperatura ambiente
 Thr: treonina
 T_m : temperatura de desnaturalización (*melting temperature*) (valor de T a mitad de la transición de desnaturalización térmica de una proteína)
 TOCSY: espectroscopía de correlación total (*total correlation spectroscopy*)
 TPPI: incremento de la fase proporcional al tiempo (*time-proportional phase incrementation*)
 tris-HCl: tris(hidroximetil)aminometano-HCl
 Trp: triptófano

TSP: ácido trimetilsilil propanoico

TTF-1: factor de transcripción tiroideo 1

(*thyroid transcription factor 1*)

Tyr: tirosina

U: unidades

UCM: Universidad Complutense de Madrid

UCSF: University of California, San Francisco

US: United States

V: valina

Val: valina

vnd: defectivo en el sistema nervioso ventral

(*ventral nervous system defective*)

W: triptófano

wt: salvaje (*wild type*)

Y: tirosina

δ : desplazamiento químico

$\Delta C_{p,b}$: diferencia de capacidad calorífica
asociada a la unión (*binding*)

ΔC_p : diferencia de capacidad calorífica entre el
estado nativo y el desnaturalizado de una
proteína

ΔG_b : variación de energía libre asociada a la
unión (*binding*)

$\Delta G_{\text{coulomb}}$: diferencia entre ΔG_{total} y $\Delta G_{\text{no polar}}$
(incluye, idealmente, sólo las
contribuciones polares a ΔG_F)

ΔG_F : variación de energía libre asociada al
plegamiento (*folding*) de una proteína

$\Delta G_{\text{no polar}}$: valor de ΔG_F a 2 M NaCl (incluye,
idealmente, sólo las contribuciones no
polares a ΔG_F)

ΔG_{total} : valor de ΔG_F a 0 M NaCl (incluye las
contribuciones polares y no polares a ΔG_F)

ΔH_b : variación de entalpía asociada a la unión
(*binding*)

ΔH_F : variación de entalpía asociada al proceso
de plegamiento (*folding*) de una proteína

ΔS_b : variación de entropía asociada a la unión
(*binding*)

ΔS_F : variación de entropía asociada al proceso
de plegamiento (*folding*) de una proteína

AGRADECIMIENTOS

Deseo expresar mi más profunda gratitud a todas aquellas personas que con sus incansables esfuerzos han aportado generosamente algo de sí mismos, ya sea creativa, física o espiritualmente, en la confección de este fragmento de ciencia.

Muy especialmente al Dr. Juan Luis Asensio, que con tanta sencillez me aceptó en su grupo y con tanto interés me facilitó iniciar un doctorado. Gracias por tu inspirador talento, tu contagioso entusiasmo, tu amabilidad y tu capacidad de esfuerzo, que han sido la piedra clave de este trabajo. Y también por la confianza y libertad que me has dado.

A la Dra. Ágatha Bastida y al Dr. Eduardo García-Junceda, por su atenta y divertida ayuda con la biología molecular, así como por su compañía y consejos diarios. También al Prof. Alfonso Fernández-Mayoralas, a la Dra. Cristina Vicent y al Dr. José Luis Chiara por su apoyo e interesantes aportes científicos y conversaciones cotidianas.

Con mucho aprecio recuerdo al Prof. Manolo Bernabé, por ser tan agradable y haberme hecho sentir tan importante en los primeros días. También ha sido muy importante la Dra. Julia Revuelta, siempre tan buena conmigo y dispuesta a aportar lo que fuera que me pudiera servir. Y muchas gracias también al Dr. Francisco Corzana, por su experta ayuda informática en nuestro análisis de secuencias, y por ser el modelo a seguir por todos los jóvenes científicos.

A la Dra. Mercedes de Frutos, y a Izaskun y el resto del laboratorio, por su simpatía y la plena disponibilidad de sus cuidados aparatos. Al Prof. Jesús Sanz, con el que pude compartir aficiones para interrumpir el trabajo de la mejor manera posible. A la gente de Fermentaciones Industriales, por acercarnos el uso de algunos materiales y técnicas. Y a la Dra. María Teresa Molina, por tantas y tan diversas conversaciones y por su cariño.

Muchas gracias también al personal no científico del Centro de Química Orgánica. En especial a Casi, Miguel y las dos Marías, que tan bien nos recibían

cada día. A Ana y su equipo de limpiadoras, que tanto nos amenizaban las tardes más tardías. Y cómo no, a Fermín y sus compañeros, siempre felizmente disponibles y con soluciones para todo.

A la parte del grupo establecida en el CIB: al Prof. Jesús Jiménez Barbero por la constante atención a mi proyecto, y por supuesto al Prof. Javier Cañada, que tantos trucos científicos me ha enseñado. También a M^a Ángeles y Carmen, y con mucho cariño a Víctor y a Silvia, con los que formé un fantástico equipo para disfrutar al máximo de los cursos de doctorado. Y al resto del personal del CIB, por permitirnos usar diversos aparatos: al Prof. Germán Rivas, a Israel, y al Prof. Antonio Romero, quien nos aportó material y conocimiento para nuestros intentos de cristalización. Por último, gracias también al Prof. José Luis García y a la Dra. Isabel de la Mata por aportar la base de partida para mi trabajo.

A los profesionales del CNIO, donde pudimos hacer importantes experimentos de dicroísmo y RMN, durante largos días de sabor a ciencia bien hecha: Dr. Francisco Blanco, Dr. Ramón Campos-Olivas, y Dr. Jesús Prieto, así como Araceli y los demás jóvenes científicos de por allí.

Y cómo no, ha sido muy importante para mí el Instituto de Química Física Rocasolano en toda su extensión, donde prácticamente he desarrollado la mitad de mi trabajo. Gracias al Dr. Miguel Ángel Treviño, a Cristina y a Yasmina por la ayuda con el FPLC. Al Prof. Carlos González por su amena cooperación en resonancia, y a la Dra. Irene Gómez Pinto por el tiempo que me concedió en San Francisco. A la Prof.^a Margarita Menéndez y a su grupo por sus sabios consejos y aportes técnicos, y por concedernos el uso de la prensa francesa, el refractómetro y el dicrógrafo... Hablando de lo cual me asalta el grato recuerdo del Dr. Pepe Sáiz. Y por último, mi máxima admiración al Dr. Douglas V. Laurents: que el mayor experto en proteínas me recibiese siempre sonriente y dispuesto a ayudarme, ilustrarme e ilusionarme, es algo a lo que concedo mucha importancia.

A la Prof.^a Isabel Guerrero y a su grupo del CBM, donde tuve el privilegio de iniciarme en el universo de la investigación: Nicole, Javier, Carlos, Laura, José, Verónica, Carmen, y especialmente Graciela, de quien recibí valiosos consejos.

A la Dra. Cristina Martínez Acedo, a Usoa, José Carlos, Antonio y al resto de profesionales del Centro de Tecnología Repsol.

Al Dr. José de Nó Sánchez de León, antiguo Director del Departamento de Postgrado y Especialización del CSIC, cuyo sentido común facilitó enormemente mi inicio de doctorado en un difícil momento.

Al Dr. José María Carazo, profesor en la Unidad de Biocomputación del Centro Nacional de Biotecnología del CSIC, por haber aceptado ser el tutor de esta tesis.

Por último, al Ministerio de Educación, Cultura y Deporte, por la concesión de mi beca/contrato y por la cordial asistencia, y al Ministerio de Economía y Competitividad por la financiación del trabajo.

Me gustaría a continuación agradecer de manera muy señalada a aquellos amigos que, además, han sido algo tan valioso como compañeros. Haber compartido tanto tiempo con ellos hace que sean totalmente innumerables los buenos recuerdos y las ocasiones en que he recibido suplementos de energía sólo con su presencia.

Una vez más, a Juan Luis, Paco, Julia y Ágatha. Muchas gracias Ágatha por una amistad de lo más especial y por honrarme con momentos tan apacibles con tu familia, a la que extendo este agradecimiento.

Al grupo de Eduardo, seguramente el laboratorio mejor organizado del mundo, del que he aprovechado numerosos materiales, aparatos y enseñanzas, y del que he respirado el mejor ambiente. Gracias Juanfran por tu cordialidad y por los buenos momentos, y por solucionarme algunos contratiempos. Gracias Montse por tu alegría, tu exquisita educación y tu original sentido artístico, y por los numerosos detalles que has tenido conmigo. Y gracias Laura por tu simpatía, encanto y bondad, y por mantener el apego con todos nosotros (por supuesto, gracias también a Kike). Del mismo modo destaca en mi recuerdo Fernando, siempre fuente de disfrute, cortesía y generosidad, e Isa Oroz, que me ha dado toda su ayuda en las etapas finales y que consigue mantener, con la ayuda de Iván, el espíritu del laboratorio. Y muchas gracias también a los entusiasmados estudiantes de Bioquímica que renovaban cada año el ya de por sí divertido día a día: Rodolfo, Ainhoa, Rocío y Leyre. Y por supuesto, muchas gracias y mucha admiración a Israel, por muchas cosas pero sin olvidar cuando cubrió mi trabajo una vez que lo necesité, o su excelente tesis, de la que he sacado valiosas ideas de formato: te considero insuperable como persona y como científico.

Al resto de amigos del departamento, que originaban un clima de trabajo inigualable. Asomarte al pasillo y ver a cualquiera de las siguientes personas era algo que animaba a cualquiera.

A Isabel, por estar siempre ahí y ayudarme siempre, por su inteligencia, y por ser siempre protagonista de todo lo bueno que se organizaba; y también a su gente y a Alberto. A Elisa, por su optimismo y sentido práctico, por su dulzura y por ser perfecta como amiga, y a Raúl, fuente de apoyo y útiles comentarios. A Guille, que seguramente tenga el récord mundial de agradecimientos en tesis doctorales, por su ingenio, gentileza, y sobre todo por su plena disponibilidad para ayudarnos a todos. A Félix, por su jovialidad, su eficiencia y

por su inestimable ayuda especialmente con mis preparativos para la estancia en San Francisco. A Juan Antonio, por su grata compañía en tantos lugares y a todas las horas del día y de la noche. Muchas gracias también a Carmen, que daba un toque de color al laboratorio y me alegraba el día cuando coincidíamos algún fin de semana. Tampoco puedo olvidarme de compañeros como Andrés, Carmen H. o Natalie, ni de la recién llegada Esther, a la que le espera lo mejor. Y muchas gracias también a Raquel, por su ternura, por ser tan atenta, y por ser una trabajadora tan ejemplar de la que siempre poder aprender.

A Ángela y Esther, piezas básicas del paraíso que era el departamento en mis primeros años. A Pablo y Tatiana por su afecto. A Florence, por su simpatía. A Bea, por su afabilidad y calidez. A Conchi, origen y destino de tantas fiestas tan importantes para esta tesis. A Samadi, por sus sabios y expertos consejos. También a Virginie, Marieta, Rafa y Elma. Y muchas gracias por supuesto a David, una persona totalmente única y que aportaba siempre una visión original de las cosas. A Igor, excelente trabajador y compañero, y la luz que brillaba durante todas las largas noches de ciencia. Y cómo no, a Jaime, abierto, divertido y noble como nadie, del que conservo fantásticos recuerdos de alegría y de fructíferas conversaciones.

A los que se fueron al poco de llegar yo, pero de los que me dio tiempo a aprender y disfrutar mucho: Paloma, Caroline y su niño Hugo, Alicia y Jason. Y en especial a Eva, con la que congenié de maravilla, y a Nuria, graciosísima y talentosa amiga, todo un modelo a seguir.

A Lourdes, por los buenos sentimientos que me inspira, por su franqueza, y por su inagotable bodega de buenos consejos. Y también por convertir Pittsburgh en una ciudad mítica para siempre, y a Cathy en su ciudadana más querida.

A Juan Carlos, por su incomparable generosidad de todo tipo, y por las fantásticas palabras que siempre guarda para mí y que blindan mi autoestima. Gracias también por tu ejemplar calidad y esmero en todo lo que haces. Y una grata mención también a tus amigos, a tu familia y a Iris.

A Tania, con la que he tenido deliciosas conversaciones y pintorescos episodios, y que ha sido una compañera de lujo en todas las etapas de la tesis y en el día a día. De igual forma, un agradecido gesto de admiración a Antonio.

Y el último lugar, que sobre todo en ciencia tiene tanta importancia, lo quiero emplear para Anita, por hacerme pasar tantos buenos momentos y tan solo buenos momentos. Lo mejor de la tesis ha sido todo lo que hemos compartido. Y muchas gracias también a Óliver y a Guillermo, su orgulloso familia, que me tratan de maravilla.

No puedo dejar de remarcar a los buenos amigos de distintos departamentos e institutos del edificio, que a lo largo de tantos años también han sido

importantes: Nacho y Miguel Ángel, tan divertidos y siempre tan atentos conmigo; Eduardo, Mercedes U., Dani, Nadia, Cristina, Maite; también Mercedes P., Fernando, Belén, Clara, Felipe, Maika, Dani, Stefi, y tantos otros en realidad... Y por supuesto, gracias a Mari Carmen Tabera, cariñosa, sensata y divertida; a Fede, impecable amigo de todos; a Jorge, por su exclusivo aporte intelectual, humorístico y social sobre todo en mis inicios; y a Laura y a Mónica, maravillosas amigas, por tantos y tan variados momentos de confianza y de diversión.

It is also very important for me to thank all the people at the University of California, San Francisco, where I was gifted to stay for four months, time enough to guess what the secrets of the United States success are. To Prof. Ken A. Dill for happily fulfilling my dream of working in that city, and for showing what a big research group director should look like. Thank you Dr. David Mobley as well, the most efficient worker I have ever met, for teaching me and recognizing my value. Also, thank you John Chodera for your intelligence and attention to my project. And of course, I owe a great deal to Scott, Banu and Kings, three good friends who helped me to feel like a true San Franciscan, and who enlightened me through their own research. I also want to thank JoAnne Williams, Vincent, Larry, Gabriella, Laura and the amazing Barbara Raymond. Far from the UCSF, I appreciate the altruistic help from Arlena and the wise advice from Prof. Eva Nogales. Finally, and most importantly, I am very grateful to Laura Wysocki, for being a key element in my learning of English, for her assistance with all the paperwork required for my stay in the US, and for her exclusive friendship that I have enjoyed in many different cities and places.

Como último apartado en estos agradecimientos, considero primordial mencionar a mis amigos y a mi familia, cuya contribución científica a esta tesis es necesariamente limitada, pero de los que proceden las aportaciones más ricas en motivación, que al final es la energía imprescindible para cualquier movimiento humano.

Los años, la cercanía total y las infinitas aventuras, que han cristalizado en una amistad incondicional, me llevan a agradecer en primer lugar a mis amigos de la carrera de Biología. En Ana, Antonio, Javi, Marta M., el glorioso Manu y mi querida Marta A. no sólo he encontrado a varios de mis mejores amigos, sino ejemplos óptimos de todo lo que puede significar un doctorado. Mención especial a la ayuda encontrada en las tesis de Antonio, Marta M. y especialmente de Ana. Gracias también a Jorge, Ude, Mónica, Willy, Salva, Carlos, Eva y, por supuesto, Kike, Miguel y la maravillosa Rocío, por lo bien que me hacéis sentir todos. Y un grato recuerdo también a las familias y amigos de todos ellos,

a muchos de los cuales tengo el honor de conocer, y que han ayudado a enriquecer mi personalidad.

También quiero recordar a mis amigos de la carrera de Bioquímica, con quienes formamos un ambiente único de compañerismo, envidia de muchas otras carreras, y tan instructivo como las materias que aprendíamos. En especial a Álvaro, Esther, Desi y Lucía.

Son también muy importantes mis amigos del colegio, que permanecen tras tantos acontecimientos y caminos dispares: María y Alberto, César y Esther, Elena y Sergio, Lola, Carlos y Paloma, y muchos otros que participaron en los mejores años de mi vida. Por supuesto, también a Víctor (y a su mujer Elena), siempre pendiente de mis avances y que nunca me ha fallado, y que conserva su original y educativo estilo. Gracias también a las familias y amigos de todos ellos, y a los ya numerosos niños. Y por último, mi más profunda gratitud a mi “amigo eterno” Gerardo, por su contagioso espíritu de superación y su formidable apoyo. Y muchas gracias también a su extraordinario hermano Florián por la escritura del prólogo de esta tesis, a su madre Loli por su cariño, confianza y criterio, y a su padre Florián, un hombre admirable a quien le habría encantado ver este trabajo completado.

A mis compañeros y amigos de la cafetería del Hipercor de Campo de las Naciones, que me han impresionado con su enorme capacidad de esfuerzo y generando un entorno de trabajo envidiable. Muy agradecidamente debo destacar a Javi E., Antonio, Javi S.E., Fran J, Álvaro B., Rodri, Roberto, Alberto, Ovidio, Dani y Raquel. Y con especial cariño a Irene, por su estimulante compañía, y cómo no a Álvaro, por su fraternal amistad y sus dosis de juventud y motivación, y a su familia, a Ulises y a Alberto. También a la gente del resto del centro, por incontables momentos de afecto y concordia.

I also have to send a very special and loud “thank you” to the funniest girls on earth: Jeanette, Christine and Susanne. Thanks for several of my most treasured moments of this period.

También a Patricia, por su afecto y por haber sido la mejor forma de interrumpir la escritura en cualquier punto. Y gracias también al apoyo de sus padres y de Night Light.

A Johanna, por ser el mejor estímulo para mi mente, el centro de mis ilusiones, y por hacerme sentir la persona más privilegiada del mundo. Por supuesto, gracias también a su cariñosa familia.

A los amigos de mi hermano, por lo agradables que son siempre conmigo y por estar presentes en multitud de momentos memorables. Especialmente a Tala, Rubén, David, Iván, Miguel y Jose.

Y por supuesto, una mención especial a esos grandes amigos, incluso hermanos, que tanto se han interesado por los progresos de mi tesis, y con los que hemos paseado momentos de pura familiaridad y humor por todos los

husos horarios de la Tierra: Javi, Christian, Rafa y Germán, y también Cris, Alicia, Almudena y Míriam. Gracias por ese importante toque no científico y por vuestra presencia diaria. Y la máxima gratitud a vuestros hermanos, padres, sobrinos y, cómo no, al joven Adrián.

A mis abuelos: Juan, Antonia, Gabriel, y sobre todo mi abuela María, que fue la mejor compañía posible cuando inicié la escritura de la tesis. También a la Tita, a Lino, a la tía Concha y al tío Antonio.

A mis tíos: Juan y Pepi, Carlos y Petri, Antonio y Olivia, Ramón e Isabel. Margot y Paco, Linito y Goyi, Mari Geni y Julio.

A mis primos, por su aliento y sus certeros comentarios: con mucho cariño a Bea (y David) y sus amigos; Esther y Ryan; Claudio; Sergio, Gemma y Jimena; María y Paco; al magnífico Jorge, Marta y Álvaro; Óscar y Sara; Cris, Isra y Patricia; Javi; Arancha; María. A Carlos, por su especial esfuerzo en la confección de la cubierta de esta tesis. A Pedro y Alicia; Cecilia; Gabi. A Juan, que me consiguió la fotografía de la cubierta, Ana y Marcos. Y a Ramón y Bea, María Isabel, y Lourdes, Curro, Javier y Jaime.

No quiero dejar de mencionar a Vicente y a siete amigos de la familia y sus aportes puntuales a este trabajo: Loli, Milagros, Mari Carmen, Tere, Mari Jose, Isabel y el Dr. Enrique González.

Finalmente y sobre todo, deseo agradecer profundamente a mi familia por el apoyo y la comprensión, por haberme regalado el tiempo que necesitaba, por la enorme generosidad de todo tipo y por haberse volcado con esta tesis: a mi padre, siempre dispuesto a aportar su tiempo y todo lo que haga falta para ayudarme, y capaz de convertir los fines de semana en puros oasis de diversión y relax. A mi hermano Germán, que se encarga de que no me tenga que preocupar por nada y de recordarme lo importante que soy, y que consigue que hasta los días de máxima concentración acaben en risas. Y también a Míriam, por su aporte de felicidad. A mi hermana Susana, por ser la más disponible y certera asesora en temas científicos, y por compartir mis ilusiones, que nadie mejor que ella sabe que se extienden más allá de este trabajo. Y a mi madre, que con lo aficionada que es a la lectura, me honra saber que éste es el libro cuyo lanzamiento más ha esperado. Gracias por tu compañía diaria, por los continuos ánimos y buenos deseos, por tus instructivas frases y consejos, y por la prioridad absoluta que me has dado. Y por si todo esto no fuese suficiente motivación, por prometerme que te leerías la tesis desde la primera hasta ésta, la última página.

